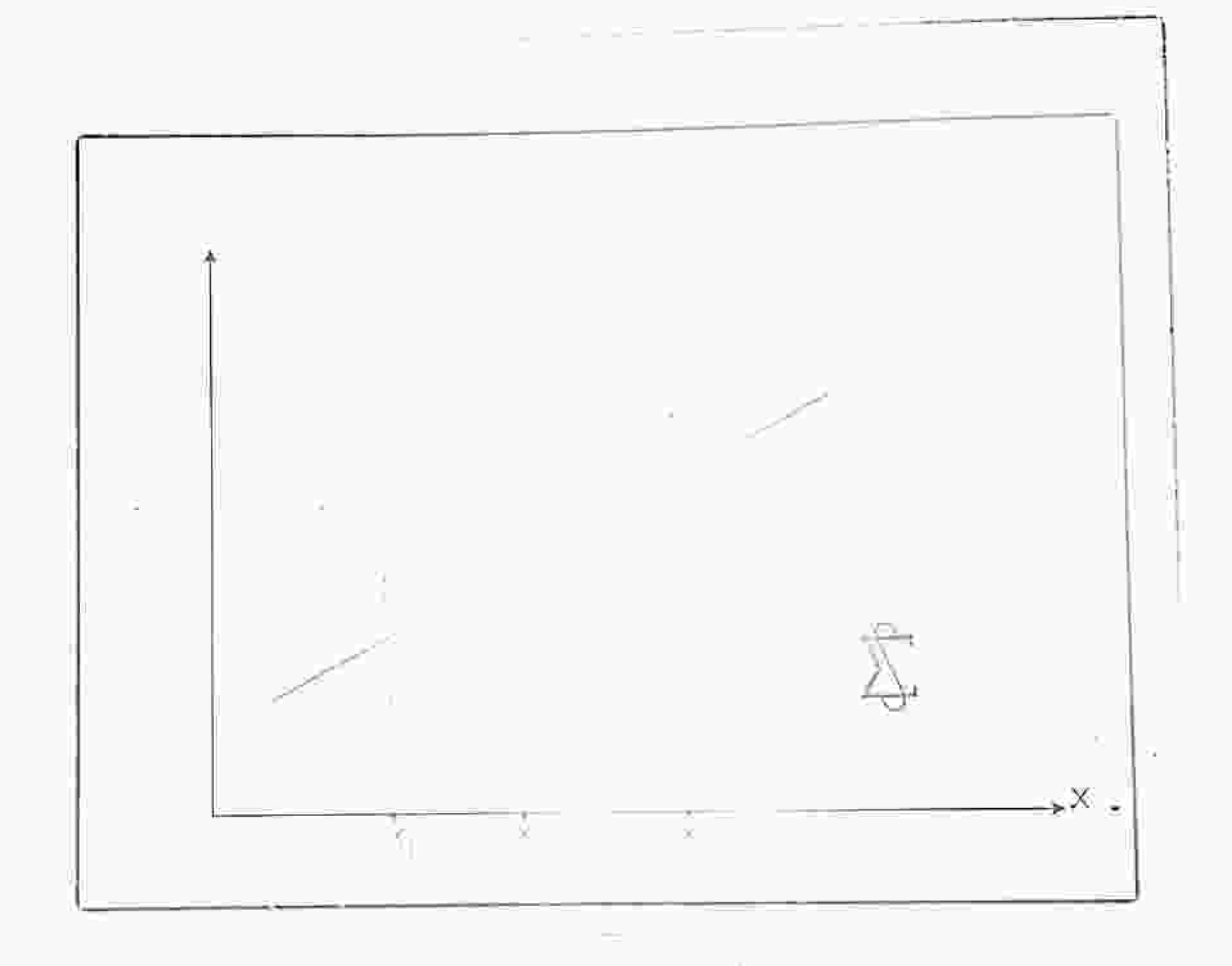
الاحصاء والاحتمالات
النظرية والتطبيق



تأليف

د. علي عبد السلام العماري
أستاذ - قسم الاحصاء
كلية العلوم - جامعة طرابلس

د. علي حسين العجيلي
أستاذ - قسم الاحصاء
كلية العلوم - جامعة طرابلس



المحتويات

الفصل الاول : عرض البيانات الاحصائية وتحليلها

الفصل الثاني : الاحتمالات

الفصل الثالث : متغيرات عشوائية في بعد واحد

الفصل الرابع : التوزيعات الإحتمالية المشتركة

الفصل الخامس : توزيعات خاصة

الفصل السادس : توزيعات المعاينة

الفصل السابع : نظرية التقدير

الفصل الثامن : اختبارات الفروض

الفصل التاسع : الإنحدار والإرتباط

الفصل العاشر : تحليل التباين

الفصل الحادي عشر : الاحصاء اللامعلمي

الفصـــــل الأول

عرض البيانات الإحصائية ووصفها وتحليلها

PRESENTATION , DESCRIPTION AND ANALYSIS OF STATISTICAL DATA

1 - I مقدمـــــــــــة Introduction

إن البيانات الاولية التي يتم جمعها عن ظاهرة معينة لا يمكن الاستفادة منها مالم تنسق حتى يمكن الالمام بما تضمنته ، وفي هذا الفصل سوف نتعرض إلى بعض الأسس المتبعة في تجميع البيانات من حيث التصنيف والتبويب ، ثم للوسائل التى تعطي فكرة واضحة وسريعة عن الظاهرة قيد الدراسة (التمثيل البياني) ، وأخيراً لبعض المقاييس الإحصائية العددية وهى تشمل مقاييس النزعة المركزية (مقاييس الموقع) ومقاييس التشتت .

1 - 2 بعض المفاهيم الإحصائية العامة Some General Concepts

تعريف (1) : علم الإحصاء

هو العلم الذي يبحث في طرائق جمع البيانات الخاصة بمختلف الظواهر وتبويبها وتحليلها للوصول إلى نتائج تساعد في اتخاذ القرارات المناسبة عندما تسود ظروف عدم التأكد ، وعلى أساس هذا التعريف يمكن تقسيم هذا المجال من المعرفة إلى نوعين هما :

الإحصاء الوصفي : (Descriptive Statistics) وهو يختص بطرائق جمع ووصف وتلخيص البيانات وذلك باستخدام الجداول التكرارية والرسومات البيانية وبعض المقاييس الإحصائية ، والإحصاء الاستنتاجي أو الاستدلالي : (Inferential Statistical) وهو يختص باستنتاج واتخاذ القرارات المناسبة للظاهرة قيد الدراسة مع حساب درجة الثقة المصاحبة لتلك القرارات والاستنتاجات .

والمهتم بهذا النوع من فروع المعرفة يجد أن لهذا العلم تطبيقات عديدة في مختلف العلوم التطبيقية والإنسانية كالعلوم الطبية ، العلوم الهندسية ، العلوم الزراعية ، العلوم الإدارية والمالية ، العلوم التربوية ، علم الاجتماع ... الخ .

تعريف (2) : المجتمع الإحصائي Statistical Population

هو مجموعة من العناصر المشتركة في الصفة التي تهم الباحث ، فمثلاً عند دراسة ظاهرة الحوادث بمدينة ما فإن المجتمع الإحصائي هو جميع الأفراد الذين تعرضوا أو ارتكبوا حوادث في تلك المدينة . أو عند دراسة مدى فعالية دواء معين لعلاج مرض ما فإن المجتمع الإحصائي هنا يشمل جميع المرضى الذين استخدموا ذلك العلاج ومصابون بذلك المرض ، وبالمثل إذا كان المطلوب معرفة مساهمة قطاع السياحة في الدخل القومي لبلد ما فإن المجتمع الإحصائي هنا يتضمن الإيرادات المتحصل عليها من جميع الفنادق والأماكن السياحية بذلك البلد .وهكذا يتم تعريف المجتمع الإحصائي على حسب الظاهرة أو المشكلة قيد الدراسة ، وغالباً ما يكون المجتمع الإحصائي مجتمعاً كبيراً وبالتالي دراسة جميع مفرداته قد يكون أمر غير متيسر وعليه نلجاء لدراسة جزء من مفرداته يطلق عليه تسمية عينة .

تعريف (3) : العينة Sample

هي مجموعة جزئية من مفردات المجتمع الإحصائي .وقد جرت العادة على اختيار مفرداتها بحيث يكون لكل مفردة من مفردات المجتمع نفس الفرصة بأن تكون من ضمن مفردات العينة وفي هذه الحالة يطلق عليها تسمية عينة عشوائية بسيطة .

تعريف (4) : المعلمة Parameter

هي قيمة عددية تستخدم لوصف خاصية معينة في المجتمع الإحصائي ، وغالباً ما تكون مجهولة .

فمثلاً متوسط دخل الفرد في بلد ما يعتبر معلمة وذلك لأنه يعكس المستوى المعيشي لأفراد ذلك البلد ، نسبة المصابين بالأمراض المعدية في بلد ما تعد معلمة وذلك لأنها تعبر عن مدى انتشار الوعي الصحي بذلك البلد ، متوسط الدخل اليومي بأحد العيادات يعتبر معلمة وذلك لأنه يعبر عن مدى إقبال المرضى على تلك العيادة ... الخ .وقد جرت العادة على أن يرمز للمعلمة بأحد الحروف اللاتينية فمثلاً متوسط المجتمع الإحصائي يرمز له بالرمز μ ، وتباينه بالرمز σ^2 ونسبة عناصره التي تتصف بخاصية معينة بالرمز P .

تعريف (5) : الإحصاءه Statistic

أي صيغة رياضية لا تعتمد على معلمات المجتمع وتستخدم لوصف خاصية معينة في العينة يطلق عليها تسمية إحصاءه . ويرمز للصيغة الرياضية لحساب متوسط العينة بالرمز $\bar{x}$ أو $\bar{y}$ ... الخ . ولتباينها بالرمز s^2 ، ولنسبة عناصرها أو مفرداتها التي قيمتها تتصف بخاصية معينة بالرمز $\hat{p}$. وكما سنرى فيما بعد في الاستدلال الإحصائي يتم استخدام متوسط وتباين ونسبة العينة في تقدير معلمات المجتمع المناظرة لها وذلك لأنه وكما أشرنا غالباً ما تكون هذه المعلمات مجهولة ويجب تقديرها .

تعريف (6) : الظاهرة Phenomena

هي صفة لعناصر تختلف من عنصر لآخر في الشكل أو النوع أو الكمية ويطلق على الصفة تحت الدراسة متغير (Variable) مثل الزمن الذي تستغرقه عملية جراحية ، طول شخص ما ، عدد السيارات المارة بإشارة ضوئية خلال فترة زمنية معينة ، ... الخ .

1 - 2 - 1 مصادر البيانات الإحصائية

هناك عدة مصادر للحصول على البيانات الإحصائية تختلف باختلاف موضوع الدراسة والغرض منها ومن أهمها :

(1) التجارب : Experiments

تعد من المصادر الهامه في الحصول على البيانات الإحصائية . وقد تكون التجربة معملية أو حقليه وذلك كما هو الحال في العلوم التطبيقية بمجالاتها المختلفة ، أو خارج المعامل كما في العلوم الإنسانيه وكذلك العلوم الإداريه والاقتصادية .

(2) الدوريات العلميه والنشرات والسجلات :

في كل مجالات المعرفه تصدر مجلات علميه بطريقة دورية سنوية أو نصف سنوية ... الخ . كما تقوم المؤسسات العامه والجهات الرسميه بإصدار نشرات تتضمن معلومات عن أنشطتها المختلفه . إضافة الى ذلك تقوم بعض الجهات الرسميه بتسجيل بياناتها في سجلات رسميه ، مثل سجلات المواليد والوفيات والطلاق والزواج والمحاكم ... الخ .

(3) الاستبيان : Questionnaires

هو عبارة عن استمارة إحصائية تحتوى على مجموعة من الأسئلة تؤدى الإجابة عليها إلى الحصول على البيانات المطلوبة .

(4) التعددات العامة : Census

في معظم دول العالم توجد مؤسسات على غرار مصلحة الإحصاءه والتعداد في ليبيا تقوم بتعدادات عامة الغرض منها حصر إمكانياتها المختلفة البشرية والزراعية والاقتصادية وذلك للحصول على بيانات تستخدم نتائجها في التخطيط للشؤون المختلفة لنشاطات الدولة . وتقوم معظم الدول بتلك التعدادات كل 10 سنوات وذلك لأنها تحتاج إلى تكاليف مادية وبشرية كبيرة .

1 - 2 - 2 أنواع البيانات الإحصائية

البيانات الإحصائية بصفة عامة يمكن تقسيمها إلى قسمين بيانات كمية (عددية) Quantitative وبيانات كيفية أو وصفية (نوعية) Qualitative .

أ - البيانات الوصفية (النوعية) Qualitative Data

وهى التي يتم تصنيف مفرداتها وفقاً لخاصية معينة في تلك البيانات فمثلاً، تصنف الإنتاج لمصنع معين من حيث المطابقة للمواصفات المطلوبة أو عدم المطابقة أو تصنيف الطلاب حسب تقديراتهم ، وقد تكون البيانات قابلة للترتيب مثل تقديرات الطلاب أو المستوى الاقتصادي ... الخ أو قد تكون غير قابلة للترتيب مثل الجنس ، وأنواع الأمراض ... الخ . فالبيانات النوعية هي بيانات عن ظواهر لا يمكن التعبير عنها عددياً حيث تكون الظاهرة قيد الدراسة مقسمة إلى صفات أو أنواع أو أزمنة .

ب - البيانات الكمية Quantitative Data

وهى التي تكون مفرداتها مقاسه بمقياس كمي وقد تكون هذه البيانات منفصلة مثل عدد . الطلاب في مراحل التعليم المختلفة أو عدد النزلاء بأحد المستشفيات بالأقسام المختلفة ... الخ . وقد تكون متصلة مثل الأطوال والأوزان ودرجات الحرارة ... الخ .

فالبيانات الكمية هي بيانات عن ظواهر يمكن التعبير عنها عددياً وهي تنقسم إلى قسمين هما منفصلة ومتصلة ، وعليه فإن الظواهر التي يمكن عدها فهي ظواهر كمية منفصلة بينما الظواهر التي يمكن قياسها فهي ظواهر كمية متصلة .

1 - 2 - 3 طرائق جمع البيانات الإحصائية

عند القيام بدراسة إحصائية لظاهرة معينة يتطلب الأمر جمع بيانات (Data) ومعلومات عن مفردات أو عناصر (وحدات) المجتمع قيد الدراسة ويتم ذلك باستخدام طريقة الحصر الشامل أو طريقة العينات .

أ - طريقة الحصر (المسح) الشامل Census

عند أتباع هذه الطريقة يتم تجميع البيانات من كل عنصر من عناصر المجتمع فإذا كنا بصدد دراسة مستوى التحصيل العلمي في جامعة العرب الطبية مثلاً يتم تجميع البيانات من كل طلبة وطالبات هذه الجامعة .وتستخدم هذه الطريقة غالباً في الحالات التالية :

(1) إذا كان المجتمع قيد الدراسة صغيراً .

(2) إذا كان المطلوب الحصول على بيانات على مستوى عالي من الدقة كما هو الحال في التعدادات العامة سواء كانت سكانية أو زراعية أو اقتصادية أو غيرها .

(3) إذا تعذر الحصول على إطار لمفردات المجتمع . فالإطار هو عبارة عن قوائم أو خرائط دالة لعناصر المجتمع قيد الدراسة .

ب - طريقة المعاينة Sampling Method

إذا تعذر استخدام طريقة المسح الشامل في الحصول على البيانات الإحصائية لأسباب عملية أو اقتصادية يتم اختيار جزء (عينة) من عناصر المجتمع قيد الدراسة بأسلوب علمي سليم ، وبتحليل بيانات العينة إحصائياً يمكن تعميم نتائجها على المجتمع ككل ، مع ملاحظة أن نتائج العينة المختارة تكون قريبة من حقائق المجتمع كلما زاد حجم العينة وكلما تم إتباع الأسلوب العلمي السليم في اختيارها وكما أشرنا سلفاً أن بعض الدراسات لا يمكن القيام بها باستخدام طريقة المسح الشامل ومن أمثلة ذلك تحليل دم مريض ، مدى مطابقة ما تم إنتاجه من قبل مصنع لبدة للأسمنت للمواصفات الليبية العالمية خلال وردية معينة ، مدى صلاحية مساحة شاسعة من الأرض لمحصول معين ... الخ .

بالإضافة إلى عدم إمكانية الإحاطة بعناصر (وحدات) المجتمع قيد الدراسة أحياناً وإلى استحالة إتباع طريقة المسح الشامل في بعض الأحيان كتحليل دم مريض مثلاً فإن هناك أسباب أخرى تدعوا إلى إتباع طريقة العينات والتي يمكن تلخيصها في النقاط التالية :

1 - الحد من التكاليف اللازمة لإجراء البحث .

2- الحد من الخطأ الناتج عن عدم الدقة في القياس وذلك لمحدودية مفردات المجتمع المختارة .

3- إذا كان المجتمع الإحصائي لا نهائي ، مثل متوسط أعمار الطلاب الذين التحقوا بالجامعة في الماضي والحاضر والمستقبل ، أو تقدير عدد السكان في الماضي والحاضر والمستقبل .

إن الهدف من العينة هو الوصول إلى استنتاج عن المجتمع الذي اختيرت منه ، وإن الخطوة الأولى في أخذ العينات هي تحديد حجم العينة (Sample Size) ، ثم البحث عن إطار المعاينة (Sampling Frame) الذي سنسحب منه العينة ، وبعدها نتبع إحدى إجراءات أو طرائق المعاينة التي سنأتي إلى ذكرها فيما بعد لاختيار العينة المطلوبة ، فالمعاينة (Sampling) هي الإجراء الذي يمكن بواسطته أن نستقرأ خصائص مجموعة كبيرة من العناصر (مجتمع ما) رغم أننا درسنا عدداً صغيراً نسبياً من عناصره (العينة) . وفي الحقيقة فإن اهتمامنا لا يكون قاصراً على عدد العناصر فقط وإنما على تنوعها أيضاً (Number and Kind) وهكذا فإن اهتمامنا الأول في البحوث هو أن يكون عدد عناصر العينة ونوعيتها ممثلة قدر الإمكان للمجتمع المستهدف بالدراسة (Target Population) حتى نتمكن من تعميم نتائجها على المجتمع المستهدف بثقة . إن العينة الجيدة نظرياً هي التي :

(1) توفر طرقاً لتحديد عدد عناصرها المطلوبة .

(2) تحدد فرصة (أو احتمال) لأن يكون أي عنصر من عناصر المجتمع المستهدف من ضمن عناصرها .

(3) تمكننا من تقدير الخطأ الناتج عن استخدام عناصرها بدلاً من استخدام كافة عناصر المجتمع.

(4) تمكننا من تحديد درجة الثقة في تقديرات المجتمع المعممة من نتائج العينة .

(5) تكون بسيطة بشكل كاف لتنفيذها في الواقع .

1 - 2 - 4 أنواع العينات وأساليب المعاينة Samples Types and Sampling Procedures

هناك أنواع متعددة من العينات ومن إجراءات سحبها يمكن للباحث أن يختار ما يتناسب منها مع ظروف الدراسة التي يقوم بها . وعموماً يمكن تقسيم العينات إلى نوعين رئيسيين هما :

I - العينات الاحتمالية (Probabilistic Samples) .

II - العينات غير الاحتمالية (الشخصية) Non - Probabilistic Samples .

أولاً : العينات الاحتمالية Probabilistic Samples

وهى العينات التي تسحب من المجتمع الإحصائي بحيث يكون لكل عنصر من عناصره فرصة أو احتمال معروف لأن يكون من ضمن عناصر العينة ، أي أن العينات الاحتمالية يتم اختيارها دون التدخل من قبل الباحث بأي شكل من الأشكال . وتمتاز العينات الاحتمالية في كونها ممثلة للمجتمع الإحصائي الذي سحبت منه بشكل جيد ، كما أنها قابلة للعديد من أساليب التحليل الإحصائي ، ويمكن تعميم نتائجها بثقة على المجتمع الإحصائي الذي تمثله . وتنقسم العينات الاحتمالية إلى خمسة أنواع رئيسية هي :

1 - العينة العشوائية البسيطة Simple Random Sample

وهى العينة التي تسحب من المجتمع بحيث يكون لكل عنصر من عناصره فرصة متساوية لأن يكون من ضمن عناصر العينة .

ولكي نحصل على عينة عشوائية فإننا ، بشكل عام ، نلجأ إلى استخدام ما يسمى بجداول الأرقام العشوائية (Tables of Random Numbers) ، مثل الجدول الموجود في نهاية هذا الكتاب . أعدت هذه الجداول بطريقة بحيث تكون فرصة اختيار أي رقم من الأرقام بين الصفر والتسعة متساوية . ولكي نستخدم جداول الأرقام العشوائية لابد لنا أن نرقم عناصر المجتمع بالأرقام من 1 إلى N حيث N تمثل عدد عناصر المجتمع . فإذا أردنا سحب عينة عشوائية تتألف من n من العناصر نحدد عدد الأعمدة (الخانات) التي سنستخدمها من جدول الأرقام العشوائية للحصول على الأرقام المطلوبة ، ونختار إحدى صفحات جدول الأرقام العشوائية بشكل عشوائي ، ثم نختار أحد الأرقام في تلك الصفحة وبطريقة عشوائية أيضاً . ونبدأ من هذا الرقم بالتحرك في الصف أو العمود الذي يقع فيه الرقم الذي تم اختياره ، مع إهمال أي عدد يتكرر أو أي عدد أكبر من حجم المجتمع . بعد ذلك نحدد عناصر المجتمع التي تحمل الأرقام التي تم اختيارها وبذلك نحصل على عينة عشوائية بسيطة . ولتوضيح طريقة استخدام جدول الأرقام العشوائية نأخذ المثال التالي :

مثال (1) : لاختيار عينة عشوائية تتألف من 5 أشخاص عند تكوين لجنة من مجتمع يتكون من 1000 شخص ، نستخدم جدول الأرقام العشوائية في نهاية هذا الكتاب بالطريقة التالية : نرقم عناصر المجتمع (الأشخاص هنا) بالأرقام من 1 إلى 1000 ثم نرجع إلى إحدى صفحات الجدول بشكل عشوائي (ولنفترض أننا أخذنا الصفحة الأولى من الجدول) ثم نختار سطر عشوائي وعمود عشوائي (وليكن السطر 5 والعمود 4 في صفحة 1 من الجدول) عندئذ سنحصل على رقم عشوائي هو 6446 ثم نستمر في نفس العمود لنحصل على الأرقام الأربعة المتبقية وهذه الأرقام هي على التوالي : 1652 ، 9451 ، 3043 ، 1406 ، ولما كان المجتمع الأصلي يحتوى على 1000 مفردة (شخص) فقط فيمكن لنا اختيار أول ثلاث مراتب في اليمين (أو اليسار) من الأرقام العشوائية التي أخذناها . وبذلك تكون اللجنة متألفة من الأشخاص الذين يحملون الأرقام التالية : 644 ، 165 ، 945 ، 304 ، 140 .

وإذا كانت جداول الأرقام العشوائية غير متوفرة لذي الباحث يمكنه أن يلجأ لطريقة ثانية من طرائق السحب العشوائي وهى طريقة صندوق القرعة (Chance Box) ، حيث يضع أوراقاً مطوية مرقمة من 1 إلى N (حيث N تمثل حجم المجتمع) داخل صندوق ما ويخلطها بشكل جيد ، ثم يسحب من الصندوق ورقة بعد ورقة ويقرأ أرقامها حتى يحدد كامل عناصر العينة.

2) العينة العشوائية الطبقية Stratified Random Sample

وتتحدد هذه العينة بتقسيم المجتمع إلى طبقات متجانسة (Homogeneous Sub - Groups) أو (Strata) ثم يسحب عدد معين من العناصر من كل طبقة بشكل عشوائي . ويتحدد العدد المسحوب من كل طبقة أما بنسبة حجم الطبقة إلى حجم المجتمع أو بالاستناد إلى العلاقة بين تباين الطبقة وتباين المجتمع . وعدد العناصر المسحوبة من كل الطبقات يؤلف حجم العينة العشوائية الطبقية . ويجب أن تكون هذه الطبقات غير متداخلة والمفردات داخل الطبقة الواحدة يجب أن تكون متجانسة ، فمثلاً من الممكن تقسيم الطلاب حسب تقديراتهم وبالتالي كل تقدير من هذه التقديرات يمثل طبقة ، أو تقسيم مجموعة من المرضى حسب الحالة المرضية .

إذاً يتضح مما سبق أن تقسيم المجتمع الإحصائي إلى طبقات ليس المقصود طبقات مادية وإنما طبقات لها علاقة بالمشكلة قيد الدراسة ، ويصلح هذا النوع من المعاينة عندما يكون المجتمع الإحصائي على درجة كبيرة من التباين (عدم التجانس) .ويتصف هذا النوع من العينات بما يلي :

أ) ارتفاع مستوى تمثيل العينة للمجتمع المستهدف بالدراسة .

ب) يمكن الحصول على نتائج جيدة الدقة من عينة حجمها صغير نسبياً .

جـ) أكثر كفاءة من غيرها في الأحوال التي تكون فيها مجموعات معينة من المجتمع ذات مواصفات خاصة هي السائدة .

د) تتطلب معرفة جيدة بالمجتمع المدروس ليتم تحديد الطبقات بشكل مناسب .

هـ) قد تتصف بشيء من التعقيد إذا كان عدد الطبقات كبيراً .

3) العينة العشوائية المنتظمة (الدورية) **Systematic Random Sample**

وهي أسهل في تطبيقها واستخدامها من العينة العشوائية البسيطة رغم أنها تعطي نتائج مشابهة لها من حيث درجة تمثيل المجتمع الإحصائي المستهدف بالدراسة وإمكانية تعميم نتائجها بثقة . ويتم اختيار أو سحب عناصر العينة المنتظمة بتحديد شيئين أساسيين : الأول هو فترة السحب (Sampling Interval) والثاني نقطة البداية (Starting Point) . وتتحدد فترة السحب بقسمة حجم المجتمع على حجم العينة . أما نقطة البداية فهي أي رقم عشوائي نختاره يكون محصوراً بين " 1 " وطول الدورة (فترة السحب) ، فمثلاً إذا كان المجتمع الإحصائي يتألف من 36 مفردة وأردنا اختيار عينة حجمها 6 فإن طول الدورة = $\frac{\text{حجم المجتمع}}{\text{حجم العينة}} = \frac{36}{6} = 6$ ،

ثم نحدد نقطة البداية أي العنصر الأول بالعينة وذلك باختيار عدداً عشوائياً من جدول الارقام العشوائية يكون محصوراً بين " 1 " و " 6 " ، والعينات التي يمكن سحبها موضحة بالجدول التالي :

العينة الأولى	العينة الثانية	العينة الثالثة	العينة الرابعة	العينة الخامسة	العينة السادسة
1	2	3	4	5	6
7	8	9	10	11	12
13	14	15	16	17	18
19	20	21	22	23	24
25	26	27	28	29	30
31	32	33	34	35	36

4) العينة العشوائية العنقودية Cluster Random Sample

وفيها يقسم المجتمع إلى عناقيد (Clusters) متنافرة بحيث يحتوى كل عنقود مختلف أنواع العناصر الموجودة في المجتمع ، ثم تسحب العينة عشوائياً من هذه العناقيد . وتجدر الإشارة هنا إلى أنه في حين يتم تقسيم المجتمع في حالة العينة الطبقية على أساس التجانس فإن التقسيم في حالة العينة العنقودية يتم على أساس تنافر العناصر لكي يكون كل عنقود ممثلاً لكامل المجتمع .

5) العينة العشوائية متعددة المراحل Multi - Stage Random Sample

تستعمل هذه الطريقة عندما يصعب الوصول مباشرة إلى كافة عناصر المجتمع المستهدف بالدراسة ، وكذلك في الأحوال التي يصعب فيها ـ نتيجة لكبر حجم المجتمع ـ إعداد إطار السحب أو تعيين تفصيلي يتضمن كافة عناصر المجتمع . وبالتالي ليس من الضروري حسب هذه الطريقة الحصول على إطار سحب كامل لعناصر المجتمع وخاصة في المرحلة الأخيرة ، وعادة ما يتم استخدام هذا النوع من المعاينة في الإحصاءات الزراعية .

أنواع الخطأ في العينات الاحتمالية :

عند القيام بدراسة ما واختيار أحد أنواع العينات الاحتمالية للحصول على بيانات ثم تحليلها وتعميم النتائج على المجتمع المستهدف بالدراسة ، قد يتعرض الباحث إلى نوعين من الخطأ وهما خطأ المعاينة وخطأ التحيز .

I) خطأ المعاينة : ينتج هذا النوع من الأخطاء بسبب وجود اختلافات وفروق بين عناصر العينة التي تم اختيارها بطريقة عشوائية وبين عناصر المجتمع المستهدف بالدراسة ، والتي شاءت الصدفة عدم اختيارها في العينة ، إلا أنه يمكن تقليل مقدار تأثير هذا النوع من الأخطاء وذلك بإتباع الطرق الإحصائية السليمة في اختيار عناصر العينة ، وكذلك بزيادة حجم العينة .

II) خطأ التحيز : هذا النوع من الخطأ يعد أكثر خطورة من خطأ المعاينة وذلك بسبب صعوبة حساب مقدار تأثيره على نتائج العينة . ويرجع الوقوع فيه لعدة عوامل أهمها :

1 ـ عدم الحصول على بيانات كاملة .

2 ـ اختيار عينة من مجتمع لا يطابق المجتمع المستهدف ، وذلك لعدة أسباب كعدم وجود إطار جيد يمكن الاعتماد عليه .

3 - استعاضة بعض عناصر العينة بعناصر أخرى من المجتمع المستهدف بالدراسة .
4 - عدم إتباع الأساليب الإحصائية السليمة في تحليل البيانات وتعميم النتائج .

ثانيا : العينات غير الاحتمالية Non - Probabilistic Samples

هناك أنواع لا حصر لها من طرائق اختيار العينات غير الاحتمالية تختلف باختلاف اتجاهات الباحثين القائمين بالدراسة . وسنتعرض هنا إلى نوعين فقط من هذه الأنواع هما :

1 - العينة العرضية (المختارة عن طريق الصدفة) Accidental Sample

وتعتمد في اختيارها على المصادفة المحضة . وتمتاز هذه الطريقة بتوفير الوقت والتكاليف ، كما يمكن من خلالها الحصول على معلومات موثوقة إذا كان المجتمع المستهدف بالدراسة على جانب كبير من التجانس (Homogeneous) ولكنها تحمل في طياتها مخاطرة التحيز (Bias) خاصة عند عدم تجانس عناصر المجتمع .

2 - عينة الحصة Quota Sample

حيث تحدد حصة مقررة لكل مجموعة أو طبقة من طبقات المجتمع المدروس ، ثم نستخدم طريقة المصادفة في اختيار مفردات العينة . وتمتاز هذه الطريقة في كونها تخفض التحيز المحتمل وقوعه في العينات غير الاحتمالية ، كما أنها تكون عملياً مفيدة في حالة عدم توفر أطر سحب العينات لطبقات المجتمع . لكنها أيضاً تحمل مخاطرة التحيز عندما لا يتحقق التوازن بين حصة الطبقات من عناصر العينة ، ومدى وزن أو أهمية هذه الطبقات في المجتمع المدروس .

حجم العينـــــة Sample Size

هناك اعتبارات أو عوامل عديدة تتحكم في اختيار حجم العينة أهمها ما يلي :

1 - التجانس (Homogeneity) : كلما ازداد تجانس (تماثل) عناصر المجتمع وقلت الفروقات بين عناصره كلما أمكن تصغير حجم العينة .

2 - اجراءات أو طريقة تحديد واختيار العينة (Sampling Procedures) حيث تؤثر نوعية العينة المختارة وطريقة اختيارها على حجم العينة .

3 - الوقت والموارد المادية والبشرية (Time , Money and Personal) المتاحة للدراسة لها أثرها الكبير في تحديد حجم العينة . وبالطبع - كقاعدة عامة كلما ازداد حجم العينة المختارة بشكل صحيح كلما ازدادت دقة وموثوقية تعميم نتائجها على المجتمع المأخوذة منه . أيضاً كلما توفر المزيد من الوقت والأموال والعناصر البشرية المشاركة في الدراسة كلما أمكن زيادة حجم العينة.

4 - درجة الخطأ المعياري المقبولة وحدود الثقة .

وأخيراً فإن فرصة اختيار نوع العينة وطريقة سحبها غالباً ما تتحدد بمدى الحاجة إلى تقسيم المجتمع إلى مجموعات أو طبقات (Strata) وبمدى توفر أطر سحب العينات (Sampling Frames) .

1 - 2 - 5 أنواع البحوث الإحصائية وخطوات القيام بها

البحوث الإحصائية يمكن تقسيمها إلى ثلاثة أنواع وهي :-

1 - البحوث الوصفية وهي التي تجمع المعلومات عن ظاهرة معينة لا لخدمة هدف بذاته محدد سلفاً ، وإنما بقصد توفير بيانات من الممكن أن تخدم أغراضاً متعددة لباحثين فيما بعد ، مثل تعدادات السكان والزراعة والصناعة والتجارة والصحة ... الخ .

2 - البحوث الإحصائية التحليلية وهي التي تجمع فيها المعلومات التي تخدم هدف معين أو تساعد في تفسير مشكلة معينة لاحظها الباحث أو لاختبار صحة فرض معين .

3 - البحوث الإحصائية التجريبية ويستخدم هذا النوع من البحوث في ميادين مختلفة كالطب والزراعة والنواحي الاجتماعية والاقتصادية .

عند القيام بأي دراسة إحصائية يجب مراعاة النقاط التالية :

1 - تحديد الهدف من الدراسة ، وبالتالي يتم تحديد البيانات التي يجب توافرها لموضوع البحث.

2 - تحديد المجتمع الإحصائي الذي ستشمله الدراسة الحظ أنه هناك نوعان من المجتمعات الإحصائية :

ا - مجتمع الهدف Target Population وهو المجتمع المستهدف بالدراسة .

ب - مجتمع العينة Sampled Population وهو المجتمع الذي سيتم اختيار العينة منه .

3 - تحديد الوحدة (المفردة) Unit ، بعد تحديد المجتمع قيد الدراسة يجب تحديد وحدات أو مفردات ذلك المجتمع تحديداً دقيقاً ، فمثلاً إذا كانت الدراسة تتعلق بالأمراض فيجب تحديد نوع المرض وإذا كانت الدراسة تتعلق بالنشاطات الحرفية فيجب تحديد نوع الحرفة ... الخ .

4 - تحديد الإطار :Frame ، بعد تحديد المجتمع وتعريف وحداته يتطلب الأمر الحصول على إطار صحيح وحديث وشامل ، فالإطار كما أشرنا في السابق هو عبارة عن خرائط أو قوائم يمكن الاهتداء بها على عناصر أو وحدات المجتمع .

5 - تحديد البيانات المطلوب جمعها بدقة تامة بحيث لا يتم جمع بيانات لا تخدم موضوع الدراسة أو يتم تجاهل بيانات غاية في الأهمية .

6 - تحديد المصادر التي ستجمع منها البيانات .

7 - تحديد طريقة جمع البيانات .

8 - تصميم الاستمارة الإحصائية :- بعد تحديد الغرض من البحث وتحديد المجتمع الذي يشمله البحث واختيار الأسلوب الذي سيتبع في تجميع البيانات الإحصائية ، يأتي بعد ذلك مرحلة تصميم الاستمارة الإحصائية إذا كانت طريقة جمع البيانات تتطلب ذلك والتي تحتوى على مجموعة من الأسئلة تؤدى الإجابة عليها إلى الحصول على البيانات المطلوبة وهناك نوعان من الاستمارات الإحصائية هما :-

أ - صحيفة الاستمارة أو الاستبيان وهى التي يقوم الشخص المستجوب بملئها بنفسه إما عن طريق الاتصال الشخصي أو عن طريق إرسالها بالبريد ويستخدم هذا النوع من الاستمارات عادة إذا كان المستوى الثقافي والتعليمي للأشخاص المبحوثين يؤهلهم للإجابة عن الأسئلة المطلوبة دون صعوبة .

ب - النوع الثاني هو كشف الأسئلة ويحتاج إلى قيام الباحث بالاتصال بالأشخاص المبحوثين عن طريق المقابلة ومساعدتهم على ملئ البيانات المطلوبة عن طريق شرح الأسئلة ومناقشة المبحوثين للتأكد من صحة البيانات ويستخدم هذا النوع عادةً في المجتمعات ذات المستوى التعليمي المنخفض وإذا كانت طبيعة البحث تتطلب جمع بيانات تحتاج إلى قيام الباحث بقياس بعض الظواهر بنفسه كقياس المساحات أو أوزان ... الخ .

وعليه فإن الاستمارة المستخدمة تتوقف على الغرض من البحث ونوع المجتمع الذي سيشمله والاستمارة الإحصائية هي الوسيلة التي تربط طالب البيانات بمصادر البيانات ، ولهذا يجب أن تكون هذه الوسيلة صالحة وقادرة على توفير البيانات المطلوبة دون خطأ أو تحيز وتشير التجارب السابقة في هذا الميدان إن فشل بعض الأبحاث الميدانية في تحقيق أهدافها يرجع غالباً إلى قصور في تصميم الاستمارة الإحصائية وسواء كانت الاستمارة المستخدمة صحيفة استبيان أو كشف أسئلة فهناك اعتبارات مختلفة يجب أن تراعى عند تصميم وصياغة الأسئلة ضماناً للحصول على بيانات دقيقة ، ومن هذه الاعتبارات :-

أ- أن تشمل الاستمارة على كل الأسئلة اللازمة للحصول على البيانات المطلوبة ، أي أنه يجب أن تكون الأسئلة متوافقة مع تحقيق أهداف دون زيادة أو نقصان ومن الأفضل صرف الوقت الكافي منذ البداية لتحديد الأهداف بشكل واضح وكذلك تحديد الملحق المطلوب ليفي بالأهداف المطلوبة وبعدها يتم التفكير بالأسئلة التي يجب أن تتضمنها الاستمارة ، إذ أن إغفال الباحث جمع بعض البيانات التي لها علاقة بالمشكلة محل البحث قد يقلل من قدرته على الاستنتاج السليم بل قد يهدر أهمية البيانات الأخرى التي جمعت .

ب- يجب أن تكون الأسئلة واضحة سهلة الفهم لا تقبل اللبس أو التأويل لأكثر من معنى منعاً للتحيز في الإجابات وهذا يتطلب وضع التعاريف المحكمة المناسبة التي تبين المقصود من كل سؤال ، وتحديد وحدات القياس المستخدمة في جمع البيانات .

جـ- يجب أن تستهدف الأسئلة الحصول على إجابات محددة واضحة ضماناً للدقة في قياس الظواهر محل الدراسة .

د- يفضل بالنسبة للأسئلة التي تستهدف قياس ظواهر غير رقمية أن لا يترك السؤال مفتوحاً بل يوضح مع السؤال الاحتمالات المتوقعة للإجابة لتفادى الإجابات الغامضة .

9 - التدريب واختبار الاستمارة الإحصائية :- من المراحل الأساسية التي يجب أن يربطها البحث الميداني إجراء برنامج تدريب للباحثين الذين سيقومون بجمع البيانات ، ويشمل هذا التدريب شرحاً وافياً لأهداف البحث ونوع المجتمع الذي سيشمله وشرحاً للتعاريف المستخدمة في الأسئلة الواردة في الاستمارة وطريقة الاتصال بالأشخاص المبحوثين وكيفية توجيه الأسئلة وكيفية الملاحظة أو القياس إذا كان الباحث سيقوم بنفسه بقياس الظواهر ... إلى آخره من الاعتبارات التي تضمن الحصول على بيانات دقيقة ، كما يفضل إجراء بحث تجريبي على عينة من المفردات لتدريب الباحثين على جمع البيانات ولاختبار الاستمارة الإحصائية والتعرف على ما قد يكون فيها من ثغرات أو مواطن قصور وتعديل الاستمارة على ضوء هذه الخبرة الميدانية.

وبما أن نجاح أي بحث ميداني يتوقف إلى حد بعيد على ما للباحثين من خبرة في مقابلة الناس وقدرتهم على أجوبة للأسئلة دون إثارة احتجاجات ، لذلك يجب على الباحث أن يكون متحلياً ببعض الصفات ليستطيع التصرف بحكمة ، ومن هذه الصفات ما يلي :-

أ - أن يكسب ثقة من يقابله ويستحق تلك الثقة .

ب - أن يخلق جواً من الود يشجع على الكلام .

جـ - أن لا يغير موضوع الحديث وأن يتجنب دور المعلم .

د – أن يطرح الأسئلة بصورة يسهل فهمها وأن يتفادى إعطاء الجواب بنفسه .

هـ – أن يتحقق من صحة الأجوبة كلما كان ذلك ممكناً وأن يعطي وقتاً كافياً للمستجوب .

و – أن يحافظ على الوقت فيحضر في الموعد المحدد .

ز – أن يتصف ببشاشة الوجه وسعة الصدر و الصبر وأن يراعي الذوق والأدب في الحديث .

ي – أن يختار الوقت المناسب لجمع البيانات وأن يراعي التقاليد في البيئة .

وأخيراً يتعين عليه عند زيارة المستجوب لأول مرة أن يشرح له بإيجاز الغرض من زيارته وإن يؤكد له أن المعلومات التي سيدلي بها ستبقى سرية للغاية بحكم القانون .

10 – التوعية : تسهيلاً لمهمة الباحث وتنويراً للرأي العام وكسباً لثقته يتعين على الجهة القائمة بالبحث أن تقوم بحملة توعية القصد منها إطلاع الرأي العام على الآتي :-

ا – الهدف من البحث وفوائده . ب – شرح طريقة القيام بالعملية .

جـ – شرح بيانات البحث والتعاريف الأساسية . د – التأكيد على سرية البيانات .

ويمكن استخدام جميع وسائل التوعية المتاحة في البلد ويتوقف اختيار وسيلة التوعية إلى حد ما على نوع البحث المراد القيام به ومن الوسائل التي يمكن استخدامها ما يلي :-

أ – مقالات في الصحف والمجلات . ب – محاضرات في المدارس والأماكن العامة .

جـ – أحاديث وندوات وإرشادات يومية بالإذاعتين المرئية والمسموعة .

د – نشرات توزع على المواطنين . هـ – لصق إعلانات على الجدران ...الخ .

11 – تحديد الزمن المناسب للدراسة .

12 – جمع البيانات وتصنيفها وعرضها بشكل يمكن الباحث من التعرف على أهم خصائصها .

13 – استخدام الأساليب الإحصائية المناسبة لتحليل البيانات من أجل الحصول على نتائج الدراسة .

1 - 3 العرض الجدولي والبياني للبيانات الإحصائية

تستخدم الجداول التكرارية والرسومات البيانية في وصف وتلخيص البيانات الإحصائية وذلك من أجل توضيح معالمها الأساسية ، وكما نعلم أن البيانات لا تفسر نفسها وبالتالي نحن بحاجة إلى طرائق تساعد تفحص هذه البيانات وتحليلها ، وإن من أهم الخدمات التي يقدمها علم الإحصاء للعلوم الأخرى هو كيفية تنظيم واختصار البيانات بشكل يمكن القارئ أو الباحث من تفهمها والوقوف على أهم خصائصها ، ومن أهم الوسائل التي يستخدمها لهذا العرض هو عمل ما يسمى بجدول التوزيع التكراري لتلك البيانات ، حيث يتم في هذا الجدول توزيع البيانات الإحصائية

المأخوذة عن ظاهرة ما على عدد معين من الفئات أو الفترات وهذا العدد تحدده ظروف الظاهرة مدار البحث .

1-3-1 العرض الجدولي

هناك عدة أنواع من الجداول التي يمكن بها وصف وتلخيص البيانات الإحصائية وذلك من أجل توضيح معالمها الأساسية بكل يسر وسهولة ، تختلف باختلاف نوع البيانات من ناحية والغرض من الدراسة من ناحية أخرى ومن أهمها :

1 - الجداول التكرارية البسيطة .

2 - الجداول التكرارية ذات الفترات .

3 - الجداول التكرارية التجميعية .

4 - الجداول التكرارية النسبية والمئوية .

5 - الجداول التكرارية المزدوجة .

(1) الجداول البسيطة

يستخدم هذا النوع من الجداول لوصف وتلخيص بيانات تتعلق بظاهرة واحدة فقط سواء كانت تلك الظاهرة كمية أو وصفية وهو أسهل وأبسط الجداول تركيباً ومفهوماً .

مثال (2) : البيانات التالية تبين فصائل الدم لعشرين مريضاً أجريت لهم عمليات جراحية في مركز طرابلس الطبي خلال أسبوع معين :

O AB O B A B O A B O A O A B O B O O AB A

والمطلوب عرض هذه البيانات في جدول توزيع تكراري .

الحـل :

فصيلة الدم	عدد المرضى (التكرار)
A	5
B	5
AB	2
O	8
المجموع	20

ألحظ أن الجدول أعلاه يطلق عليه جدول تكراري بسيط نوعي وذلك لأن المرضى تم تصنيفهم حسب نوع فصيلة الدم .

مثال (3) : البيانات التالية تمثل درجات مجموعة من الطلبة في مقرر ما ، والمطلوب توضيح المعالم الأساسية لهذه البيانات وذلك من خلال وضعها في جدول تكراري بسيط .

92 ، 50 ، 30 ، 40 ، 42 ، 60 ، 30 ، 40 ، 42 ، 50 ، 50 ، 65 ، 65 ، 60 ، 60 ،
60 ،40 ، 65 ، 50 ، 60 ، 60 ، 42 ، 50 ، 42 ، 50 ، 60 ، 65 ، 60 ، 42 ، 50 ،
65 ، 68 ، 68 ، 68 ، 65 ، 80 ، 92 ، 92 ، 65 ، 60 ، 68 ، 50 ، 65 ، 68 ، 60 ،
68 ، 92 ، 65 ، 80 ، 92

الحل :

رتب الدرجات ترتيباً تصاعدياً أو تنازلياً ثم حدد عدد الطلبة (التكرار) المناظر لكل درجة . ثم يتم تلخيص البيانات في جدول التوزيع التكراري البسيط التالي :

الدرجة	30	40	42	50	60	65	68	80	92	المجموع
عدد الطلبة (التكرار)	2	3	5	8	10	9	6	2	5	50

(2) الجداول التكرارية ذات الفترات

إذا كانت البيانات كثيرة وتكراراتها قليلة نجد أن الجدول التكراري البسيط لن يفي بالغرض المطلوب من حيث وصف وتلخيص البيانات ، وذلك لأن وضع البيانات في الجدول التكراري البسيط لن يختلف كثيراً عن وضعها الأصلي ، وفي مثل هذه الحالة يتم استخدام نوع آخر من الجداول التكرارية وهو ما يسمى بالجدول التكراري ذي الفترات حيث يتعامل هذا النوع من الجداول مع البيانات كمجموعات بدلاً من التعامل معها مفردة مفردة كما هو الحال في الجدول التكراري البسيط .

وتتلخص خطوات تكوين جدول تكراري ذي فترات في الآتي :

1 - تحديد المدى : وهو المجال الذي تنتشر فيه البيانات حيث : المدى = أكبر قيمة - أصغر قيمة .

المطلوبة لتكوين الجدول ، وسنرمز لهذا العدد بالرمز k

2 - تحديد عدد الفترات (الفئات) المطلوبة لتكوين الجدول ، وسنرمز لهذا العدد بالرمز k ويمكن الحصول على قيمة تقريبية لعدد الفترات باستخدام بعض المعادلات الرياضية التي وضعها كل من " ستيرجس " Sturges " ويول " Yole " وهما :

أ - معادلة ستيرجس : $k = 2.5\sqrt[4]{n}$ ، حيث n = عدد القيم .

ب - معادلة يول : $k = 1 + 3.322 \ Log_{10} n$.

إن الصيغتين (أ) و (ب) يمكن استخدام أي منهما كمؤشر في تحديد العدد المناسب ولكن ليس بالضرورة استخدام العدد الناتج من أي منهما ، حيث من الممكن استخدام عدد أكبر أو أصغر من ذلك وهذا أمر تقرره ظروف الظاهرة مدار البحث وكذلك وجهة نظر الباحث ، مع مراعاة ألا يقل عدد الفترات عن " 5 " ولا يزيد على " 20 " .

3 - تحديد طول الفترة : طول الفترة = $\frac{\text{المدى}}{\text{عدد الفترات}}$.

مع مراعاة تحقق المتباينة التالية : طول الفترة × عدد الفترات ≥ المدى .

وعادة ما تكون الفترات متساوية الطول إلا في حالات استثنائية التي يستحيل فيها ذلك من الناحية العلمية .

4 - تحديد بداية ونهاية كل فترة على أن تكون بداية الفترة الأولى أصغر من أو تساوى أصغر مفردة في البيانات ونهاية الفترة الأخيرة أكبر من أو تساوى أكبر مفردة في البيانات .

5 - تحديد عدد القيم (أو المفردات أو المشاهدات) التي تقع في كل فترة على أن تكون لكل قيمة فترة واحدة وواحدة فقط تنتمي إليها وهو ما يسمى بالتكرار (f) frequency .

وعند تكوين الجدول التكراري ذي الفترات ينبغي مراعاة النقاط التالية :-

أ - هناك عدة طرائق لتكوين الجدول التكراري ذي الفترات تتفق جميعها في الأسس ولكنها تختلف في طريقة العرض ، ومن أمثلة ذلك :

1 - الفترة 10 - ´20 تحتوى على كل البيانات التي أكبر من أو تساوى 10 إلى أقل من 20 .

2 - الفترة ´10 - 20 تحتوى على كل البيانات الأكبر من 10 إلى 20 (بما فيها 20) .

3 - التعامل مع الحدود الحقيقية للفترات فمثلاً الفترة 10 - 20 حدها الأدنى الظاهري يساوى 10 وحدها الأعلى الظاهري يساوى 20 بينما الحد الأدنى الحقيقي والأعلى الحقيقي لهذه الفترة على الترتيب هما 9.5 و 20.5 .

وبصفة عامة : الحد الأدنى الحقيقي = الحد الأدنى الظاهري - 0.5 وحدة قياس

الحد الأعلى الحقيقي = الحد الأعلى الحقيقي + 0.5 وحدة قياس

ب - أتباع طريقة واحدة في تحديد الحدود السفلي والعليا للفترات حتى لا يكون هناك تداخل بين الفترات .

حـ - يفضل أن تكون البيانات داخل كل فترة أقرب ما يمكن إلى منتصفها (مركزها) كل ما أمكن وذلك حتى يتسنى لنا الحصول على معلومات أكثر دقة وواقعية ، حيث :

مركز الفترة = (الحد الأدنى + الأعلى) ÷ 2 .

د - الجدول التكراري ذي الفترات قد يكون مقفل منتظم ، مفتوح من أعلى ، مفتوح من أسفل ومن أعلى ، غير منتظم . ويفضل عدم التعامل مع الجداول التكرارية ذات الفترات المفتوحة كلما أمكن ذلك .

هـ - في حالة الجداول غير المنتظمة (أطوال فتراتها غير متساوي) يجب استخدام ما يسمى بالتكرار المعدل في بعض الاحيان ، حيث :

التكرار المعدل لأي فترة = تكرار الفترة ÷ طولها .

و - الجدول التكراري ذي الفترات قد يكون متصلاً وفي مثل هذه الحالة تكون الفترات متلاصقة (متصلة) بمعنى نهاية أي فترة هي بداية للفترة التي تليها ، وإن لم يكن كذلك يتم التعامل مع الحدود الحقيقية للفترات خاصة في العرض البياني (كما سنرى فيما بعد) .

مثال (4) : البيانات التالية تمثل الأرقام الشهرية لدرجات الحرارة المئوية القصوى في ليبيا خلال فترة 1971 - 1975 م .(المصدر : أمانة التخطيط ، مصلحة الإحصاء والتعداد ، المجموعات الإحصائية 73 ، 74 - 1975 م . أي الجدول الثاني صـ3) . والمطلوب توضيح المعالم الأساسية لهذه البيانات وذلك من خلال وضعها في جدول توزيع تكراري ذي فترات .

38	36	41	36	30	28	22	16	37	31
40	39	36	40	33	31	31	15	15	20
23	41	36	32	26	19	18	23	23	30
35	28	27	21	17	19	17	30	38	30
22	23	16	15	22	29	36	32	35	37
15	15	19	34	35	36	41	40	34	30

الحـــل :

المدى = أكبر قيمة - أصغر قيمة = 41 - 15 = 26 درجة مئوية .

عن طول الفترة وعدد الفترات لذلك سيتم استخدام معادلة
حيث أنه لا توجد لدينا معلومات
ستيرجس أو معادلة يول للحصول على عدد تقريبي للفترات .

فباستخدام معادلة يول نجد أن :

$$k = 2.5\sqrt[4]{n} \quad = 2.5\sqrt[4]{60} = 2.5 \times 2.7832$$
$$= 6.958 \cong 7$$

أو باستخدام معادلة ستيرجس :

$$k = 1 + 3.322 \ \mathrm{Log}_{10}\, n$$
$$= 1 + 3.322 \ \mathrm{Log}\, 60 = 1 + 3.322 \times 1.7782$$
$$= 1 + 5.9070 = 6.907 \cong 7$$

وعليه فإن عدد الفترات المناسب يساوى 7 .

وبذلك فإن : طول الفترة = المدى ÷ عدد الفترات

3.7193 = 7 ÷ 26 =

4 ≅

وبذلك يمكن تلخيص البيانات المعطاة في جدول التوزيع التكراري التالي :

الفترات (درجة الحرارة)	15 -	19 -	23 -	27 -	31 -	35 -	39 - 43	المجموع
عدد الأشهر (التكرار)	10	7	5	9	8	14	7	60

(3) الجداول التكرارية النسبية والمئوية

بالإضافة إلى الجداول السابقة لوصف وتلخيص وتوضيح البيانات المتعلقة بالظاهرة قيد الدراسة هناك انواع أخرى من الجداول التكرارية وهى الجداول التكرارية النسبية والمئوية . وهذا النوع من الجداول له عدة استخدامات وفوائد حيث يوضح نسبة توزيع التكرار الكلى على الفترات. فالتكرار النسبي لأي فترة هو عبارة نسبة المفردات التي تنتمي لتلك الفترة ، حيث
التكرار النسبي = تكرار الفترة ÷ مجموع التكرارات .

وبضرب التكرار النسبي في 100 يتم الحصول على ما يسمى التكرار المئوي ، حيث

التكرار المئوي = التكرار النسبي × 100 .

والجدير بالملاحظة هنا مراعاة النقاط التالية :

1 - لا يمكن أن يكون التكرار العادي كسراً ، بل يجب أن يكون عدداً صحيحاً موجباً .

2 - التكرار النسبي يجب أن يكون كسراً موجباً ومجموع التكرار النسبي لجميع الفترات = 1 .

3 - مجموع التكرار المئوي لجميع الفترات = 100 .

4 - يفيد التكرار النسبي في تقليص الشكل البياني عندما يكون عدد القيم كبيراً ، بينما يفيد التكرار المئوى في إظهار الشكل البيانى عندما يكون عدد القيم صغيراً .

مثال (5) : من جدول التوزيع التكراري بالمثال السابق اوجد التوزيع التكـرارى النسـبى والمئوى لدرجات الحرارة .

الحـل :

الفترات	التكرار	التكرار النسبى	التكرار المئوى٪
15 -	10	0.167	16.7
19 -	7	0.117	11.7
23 -	5	0.083	8.3
27 -	9	0.150	15
31 -	8	0.133	13.3
35 -	14	0.233	23.3
39 - 43	7	0.117	11.7
المجموع	60	1	100

(4) الجداول التكرارية التجمعية

تستخدم الجداول التجمعية عندما نود الحصول على عدد المفردات التى تزيد أو تقل عن قيمة معينة كما تستخدم في حساب بعض المقاييس الإحصائية (كما سنرى فيما بعد) وتجدر الإشـارة هنا إلى أن هناك جداول تكرارية متجمعة صاعدة وجداول تكرارية متجمعة هابطة ، ومدينا أيضـاً بالإمكان إيجاد نوع أخر من الجداول التكرارية النسبية المتجمعة الصاعدة أو الهابطة .

مثال (6) : من بيانات المثال السابق كون كل من :

أ - جدول التكرار المتجمع الصاعد .   ب - جدول التكرار المتجمع الهابط .

جـ - جدول التكرار النسبي المتجمع الصاعد .   د - جدول التكرار النسبي المتجمع الهابط .

الحـــل :

أ - جدول التكرار المتجمع الصاعد :

درجه الحرارة	التكرار	الحدود العليا للفترات	التكرار المتجمع الصاعد
- 15	10	أقل من 19	10
- 19	7	أقل من 23	17
- 23	5	أقل من 27	22
- 27	9	أقل من 31	31
- 31	8	أقل من 35	39
- 35	14	أقل من 39	53
43 - 39	7	أقل من 43	60
المجموع	60		

ب - جدول التكرار المتجمع الهابط :

الفترات (درجه الحرارة)	التكرار (عدد الأشهر)	الحدود السفلى للفترات	التكرار المتجمع الهابط
- 15	10	15 فأكثر	60
- 19	7	19 فأكثر	50
- 23	5	23 فأكثر	43
- 27	9	27 فأكثر	38
- 31	8	31 فأكثر	29
- 35	14	35 فأكثر	21
43 - 39	7	39 فأكثر	7
المجموع	60		

جـ - جدول التكرار النسبي المتجمع الصاعد :

الفترات (درجة الحرارة)	التكرار (عدد الأشهر)	التكرار النسبي	التكرار النسبي المتجمع الصاعد
15 -	10	0.167	0.167
19 -	7	0.117	0.284
23 -	5	0.083	0.367
27 -	9	0.150	0.517
31 -	8	0.133	0.650
35 -	14	0.233	0.883
43 - 39	7	0.117	1
المجموع		1	

د - جدول التكرار النسبي المتجمع الهابط :

الفترات (درجة الحرارة)	التكرار (عدد الأشهر)	التكرار النسبي	التكرار النسبي المتجمع الهابط
15 -	10	0.167	1
19 -	7	0.117	0.833
23 -	5	0.083	0.716
27 -	9	0.150	0.633
31 -	8	0.133	0.483
35 -	14	0.233	0.350
43 - 39	7	0.117	0.117
المجموع	60	1.0	

(5) الجداول التكرارية المزدوجة

يستخدم هذا النوع من الجداول في وصف وتلخيص البيانات المتعلقة بدراسة ظاهرتين في آن واحد وقد يكون الجدول المزدوج كمي أو نوعى أو خليط (كمي ونوعى) ومن أمثلة ذلك الجداول التاليه :

أ - جدول تكراري مزدوج كمي :

الطول / الوزن	120 -	140 -	180- 160	المجموع
20 -	10	8	6	24
40 -	15	22	10	47
60 -	12	17	4	33
100 - 80	8	6	2	16
المجموع	45	53	22	120

ب - جدول تكراري مزدوج نوعى :

التدخين / الإصابة	مدخن	غير مدخن	المجموع
مصاب	300	40	340
سليم	10	250	260
المجموع	310	290	600

جـ - جدول تكراري مزدوج خليط (كمي ونوعى) :

عدد الأطفال / المستوى التعليمي لرب الأسرة	0 -	2 -	4 -	6 -	10 - 8
آمي	5	6	10	4	9
أساسي	3	8	6	12	10
متوسط	1	12	14	9	11
عالي	6	10	11	14	3

مثال (7) : بفرض أن البيانات التالية تمثل أجمالي ما أنفقه 75 شخص خلال أسبوع :

72 68 53 73 82 68 78 66 62 65 74 73 67 73
81 63 63 83 60 79 75 71 79 62 69 97 78 62
76 65 82 78 75 73 66 75 82 73 84 77 69 74
60 96 78 79 71 85 75 60 90 71 79 83 75 61
65 75 87 74 85 91 80 79 89 76 93 73 57 90
62 88 68 76 83

المطلوب :

أ – وضع هذه البيانات في جدول توزيع تكراري ذي فترات متصلة وأخر ذي فترات منفصلة ثم تحويله إلى جدول توزيع تكراري بفترات متصلة .

ب – أيجاد التكرار النسبي والتكرار المنوى .

الحـــل :

أ – لوضع هذه البيانات في جدول توزيع تكراري ذي فترات متصلة منفصلة نتبع الأتي :

1 – عدد الفترات : $k = 2.5\sqrt[4]{n} = 2.5\sqrt[4]{75}$

$= 2.5 \times 2.9428$

$= 7.4$

وبذلك يكون عدد الفترات المناسب يساوى 8 تقريباً .

2 – تحديد المدى التي تنتشر فيه البيانات حيث المدى = 97 − 53 = 44 .

3 – طول الفترة = المدى ÷ عدد الفترات = 44 ÷ 8 = 5.5 ≅ 6 تقريباً .

بعد إتمام هذه الخطوات يمكننا الأن كتابة جدول التوزيع التكراري ذي الفترات المتصلة والمنفصلة كما يلي :

1) جـــدول توزيع تكراري بفتـرات متصلة :

يقال بأن جدول التوزيع التكراري ذي فترات متصلة إذا كانت بداية كل فترة من ضمن بياناتها بينما نهايتها من ضمن بيانات الفترة التي تليها ، ولهذا السبب في بعض الأحيان لا نكتب نهاية الفترة وذلك تفادياً للبس الذي من الممكن حدوثه عند بعض القراء إلا الفترة الأخيرة بالجدول كما يتضح في الجدول الأتي :

التكرار (f_i)	الفترات
1	- 50
5	- 56
12	- 62
15	- 68
22	- 74
11	- 80
6	- 86
3	98 - 92
$\sum_{i=1}^{8} f_i = 75$	المجموع

ألحظ أنه تتم قراءة الفترات بهذا الجدول كما يلي :-

الفترة الأولى : من 50 إلى أقل من 56 وذلك لأن 56 من ضمن بيانات الفترة الثانية وليست الأولى .

الفترة الثانية : من 56 إلى أقل من 62 وذلك لأن 62 من ضمن بيانات الفترة الثالثة وليست الثانية ... وهكذا بقية الفترات .

2) جدول توزيع تكراري بفترات منفصلة :

إذا كان جدول التوزيع التكراري ذي فترات منفصلة فهذا يعني أن بداية ونهاية الفترة تنتمي لنفس الفترة .

الحدود الحقيقية للفترات	التكرار (f)	الفترات
55.5 - 49.5	1	55 - 50
61.5 - 55.5	5	61 - 56
67.5 - 61.5	12	67 - 62
73.5 - 67.5	15	73 - 68
79.5 - 73.5	22	79 - 74
85.5 - 79.5	11	85 - 80
91.5 - 85.5	6	91 - 86
97.5 - 91.5	3	97 - 92

ألحظ أنه تتم قراءة الفترات بهذا الجدول كما يلي :-

الفترة الأولى : من 50 إلى 55 وذلك لأن كلاهما من ضمن بيانات الفترة الأولى .

الفترة الثانية : من 56 إلى 61 وذلك لأن كلاهما من ضمن بيانات الفترة الثانية .

... وهكذا بقية الفترات .ولإيجاد الحدود الحقيقية للفترات إما أن نستخدم التعريف حيث كما أشرنا سابقاً :

الحد الأعلى الحقيقي = الحد الأعلى الظاهري + 0.5 وحدة قياس .

و

الحد الأدنى الحقيقي = الحد الأدنى الظاهري - 0.5 وحدة قياس .

أو نتبع الآتي :

•- الحد الأعلى الحقيقي للفترة الأولى والحد الأدنى الحقيقي للفترة الثانية = $\frac{55+56}{2}=55.5$.

•- الحد الأعلى الحقيقي للفترة الثانية والحد الأدنى الحقيقي للفترة الثالثة = $\frac{61+62}{2}=61.5$.

⋮ ⋮

•- الحد الأعلى الحقيقي للفترة السابعة والحد الأدنى الحقيقي للفترة الثامنة = $\frac{91+92}{2}=91.5$.

أما في ما يخص الحد الأعلى الحقيقي للفترة الثامنة والحد الأدنى الحقيقي للفترة الأولى يتم حسابهما كما يلي :

حيث أن: طول الفترة (من الجدول) = الحد الأعلى الحقيقي - الحد الأدنى الحقيقي ، وعليه فإن الحد الأدنى الحقيقي للفترة الأولى = 55.5 - طول الفترة = 55.5 - 6 = 49.5 .

الحد الأعلى الحقيقي للفترة الثامنة = 91.5 + طول الفترة = 91.5 + 6 = 97.5 .

إذن مما سبق يتضح أنه عند إيجاد الحدود الحقيقية للفترات يتضمن كل حد من الحدود رقم عشري واحد إضافي عن القيم المعطاة ، وبهذه الكيفية نضمن عدم وجود أي قيمة واقعة على الحد الحقيقي للفترة ومنها نتفادى آي لبس في تحديد لأي فترة تنتمي قيمة معينة .

ب - التكرار النسبي والتكرار المئوي :

الفترات	التكرار (f)	التكرار النسبي	التكرار المئوي (%)
50 -	1	0.013	1
56 -	5	0.067	7
62 -	12	0.16	16
68 -	15	0.2	20
74 -	22	0.293	29
80 -	11	0.146	15
86 -	6	0.08	8
92 -	3	0.04	4
المجموع	75	0.999	100

1-3-2 العرض البياني

بالإمكان وصف وتلخيص البيانات الإحصائية باستخدام الرسومات البيانية والأشكال الهندسية حيث الرسم لغة الشعوب . وهناك عدة أنواع من الرسومات البيانية تختلف باختلاف نوع البيانات من ناحية والغرض من الرسم من ناحية أخرى .

I - العرض البياني للبيانات الكمية

(أ) المدرج التكراري :

يستخدم المدرج التكراري لوصف وتلخيص البيانات الكمية بيانياً وذلك بعد وضعها في جدول تكراري ذي فترات حيث يتم تمثيل كل فترة بمستطيل يكون أحد أضلاعه طول الفترة والضلع

الأخر تكرارها ، ويتم تحديد الفترات على المحور الأفقي بينما التكرارات المناظرة لها على المحور الرأسي ، وعند رسم المدرج التكراري ينبغي مراعاة ما يلي :

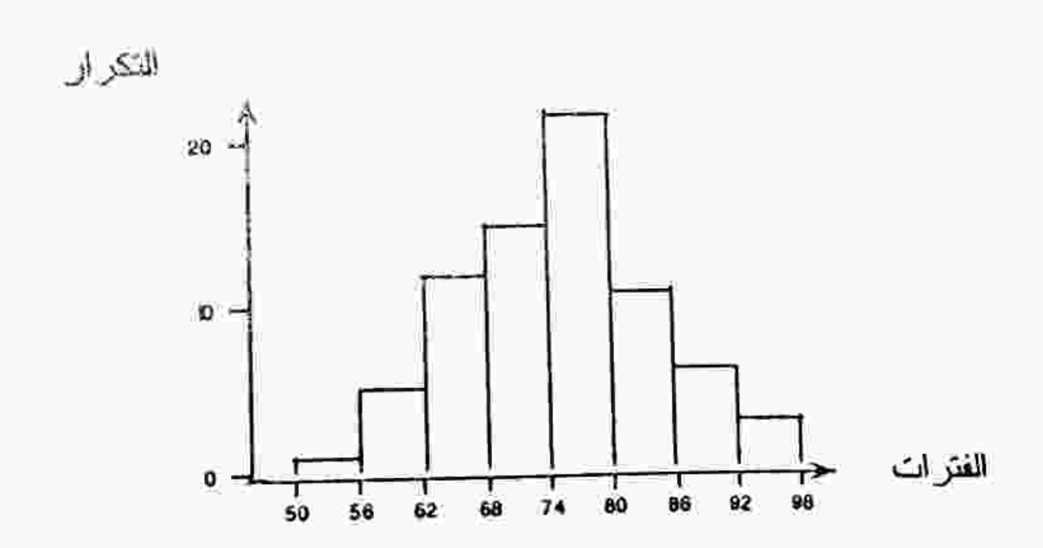
1 - أن تكون الفترات متصلة بمعنى نهاية أي فترة هي بداية للفترة التي تليها وإن لم تكن كذلك يتم التعامل مع الحدود الحقيقية للفترات .

2 - أن تكون الفترات ذات أطوال متساوية وإن لم تكن كذلك يتم استخدام التكرار المعدل الذي يتناسب مع طول الفترة .

ألحظ أن المساحة الكلية للبيانات تقع تحت المدرج التكراري وكل مستطيل من مستطيلات المدرج التكراري يمثل جزء من هذه المساحة . وتتم المقارنة بين المستطيلات المختلفة من حيث كمية المعلومات التي يتضمنها من خلال حساب مساحة كل مستطيل على حده وهذه المقارنة صحيحة إذا كانت الفترات ذات أطوال متساوية أما إذا كانت تلك الأطوال غير متساوية فإن هذه المقارنة تجر إلى استنتاج خاطئ إلا إذا تم استخدام التكرارات المعدلة .

مثال (8) : مثل بيانات المثال السابق بيانياً باستخدام المدرج التكراري .

الحـــل :



شكل (1) : المدرج التكراري

(ب) المضلع التكراري :

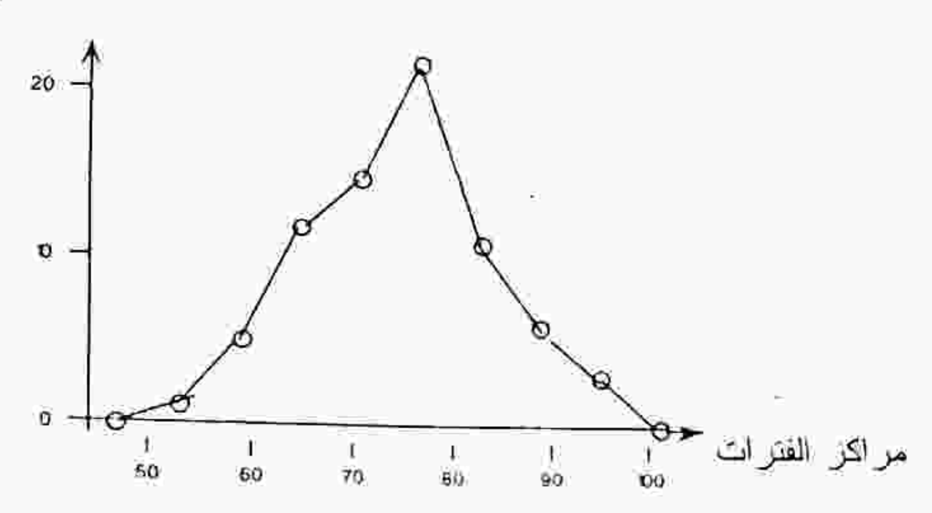
يستخدم المضلع التكراري غالباً لوصف وتلخيص البيانات الكمية بيانياً حيث يتم تحديد نقاط مراكز الفترات مع التكرارات المناظرة لها على المحورين الأفقي والرأسي بحيث تمثل كل نقطة مركز الفترة والتكرار المناظر لها ، ثم يتم توصيل تلك النقاط على الترتيب بخطوط مستقيمة نتحصل منها على ما يسمى بالمضلع التكراري وحيث أن تكرار الفترة ما قبل الأولى والفترة ما بعد الأخيرة يساوى صفر ، فإنه بالإمكان إضافة هاتين الفترتين إلى الرسم وذلك من أجل الحصول عل شكل مغلق (مضلع) .

ألحظ أنه إذا كان الغرض من الرسم وصف وتلخيص بيانات تتعلق بظاهرة كمية واحدة فقط فإنه بالإمكان استخدام المدرج التكراري أو المضلع التكراري لأن كلاهما في مثل هذه الحالة يفي بالغرض المطلوب ، أما إذا كان الغرض من الرسم مقارنة ظاهرتين كميتين أو أكثر بيانياً استخدام المضلع التكراري مع مراعاة أن تكون أطوال الفترات في المجموعات المراد مقارنتها متساوية .

مثال (9) : مثل بيانات المثال (7) بيانياً باستخدام المضلع التكراري .

الحـــل :

الفترات	التكرار (f)	مراكز الفترات
- 44	0	47
- 50	1	53
- 56	5	59
- 62	12	65
- 68	15	71
- 74	22	77
- 80	11	83
- 86	6	89
- 92	3	95
104 - 98	0	101



شكل (2) : المضلع التكراري

(جـ) المنحنى التكراري :

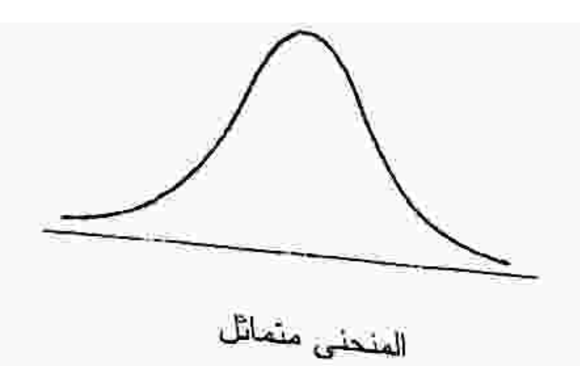
إذا كانت البيانات كثيرة وعدد الفترات كبيراً نجد أن أطوالها قصيرة وفي هذه الحالة كل من المضلع التكراري والمدرج التكراري يؤول إلى منحنى يطلق عليه تسمية المنحنى التكراري ، ويتم رسم المنحنى التكراري بنفس الكيفية التي تم بها رسم المضلع التكراري مع مراعاة النقاط التالية :

1 - يتم توصيل النقاط بخط يمهد باليد .

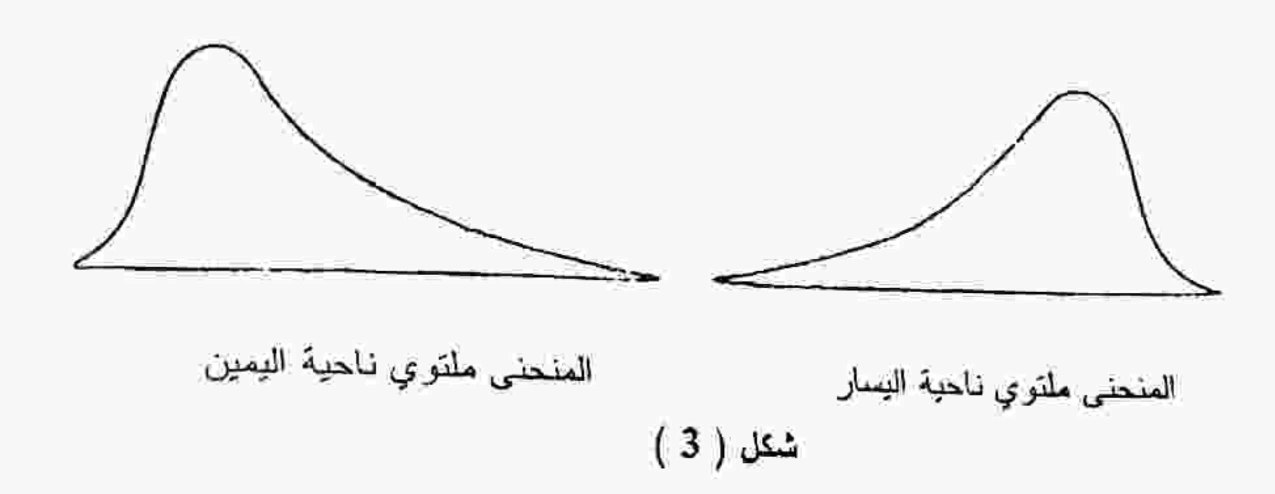
2 - ليس من الضروري أن يمر المنحنى بكل النقاط ولكن يجب أن يمر بأكبر عدد ممكن منها بحيث النقاط التي لا يمر بها يكون بعدها عنه أقل ما يمكن (وذلك من أجل الحصول على منحنى خالي من التعرجات) .

3 - لا داعي إلى إضافة فترة ما قبل الأولى وأخرى ما بعد الأخيرة من أجل الحصول على شكل مغلق .

4 - يستخدم المنحنى التكرارى عادة لوصف المجتمعات الكبيرة ، ويكون المنحنى على أشكال عديدة أشهرها المنحنى المتماثل والمنحنى الملتوى وذلك كما هو موضح في شكل (3) أدناه :



المنحنى متماثل



المنحنى ملتوي ناحية اليمين

المنحنى ملتوي ناحية اليسار

شكل (3)

(د) منحنى التكرار المتجمع الصاعد والهابط :

يتم الحصول على منحنى التكرار المتجمع الصاعد بتحديد نقاط الحدود العليا للفترات مع التكرار المتجمع الصاعد المناظر لها على المحورين الأفقي والرأسي على الترتيب وبتوصيل هذه النقاط بمنحنى نتحصل على ما يسمى بمنحنى التكرار المتجمع الصاعد . وبالمثل يتم رسم منحنى التكرار المتجمع الهابط بنفس الكيفية وذلك باستخدام الحدود السفلي للفترات مع التكرار المتجمع الهابط المناظر لها . مع ملاحظة أن يكون جدول التوزيع التكراري ذي فترات متصلة أما إذا كانت الفترات منفصلة فيجب التعامل مع الحدود الحقيقية للفترات .

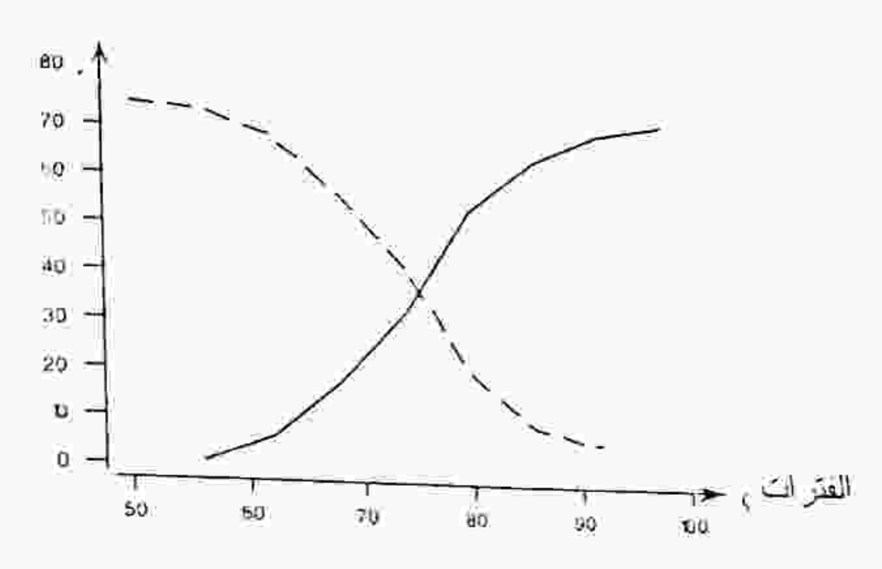
مثال (10) :

الحـــل :

الفترات	التكرار (f)	الحدود العليا للفترات	التكرار المتجمع الصاعد
56 - 50	1	أقل من 56	1
62 - 56	5	أقل من 62	6
68 - 62	12	أقل من 68	18
74 - 68	15	أقل من 74	33
80 - 74	22	أقل من 80	55
86 - 80	11	أقل من 86	66
92 - 86	6	أقل من 92	72
98 - 92	3	أقل من 98	75

بينما يتم إيجاد التكرار المتجمع الهابط كما يلي :

الحدود السفلي للفترات	التكرار المتجمع الهابط
50 فأكثر	75
56 فأكثر	74
62 فأكثر	69
68 فأكثر	57
74 فأكثر	42
80 فأكثر	20
86 فأكثر	9
92 فأكثر	6



شكل (4) : المنحنى التكراري المتجمع الصاعد والهابط

ملحوظة :-

لتمثيل البيانات باستخدام المدرج التكراري أو المنحنى التكراري المتجمع الصاعد والهابط يجب أن يكون جدول التوزيع التكراري ذي فترات متصلة وان لم يكن كذلك فيجب إيجاد الحدود الحقيقية للفترات أما إذا كان المطلوب تمثيلها بيانياً باستخدام المنحنى التكراري أو المضلع التكراري ليس بالضرورة أن يكون جدول التوزيع التكراري ذي فترات متصلة وذلك لأننا نستخدم مراكز الفترات مفترضين أن البيانات تتوزع بطريقة منتظمة بكل فترة أو أنها قريبة من مركز تلك الفترة .

مثال (11) : أوجد الحدود الحقيقية للفترات التالية :

أ- 8 - 13 ب- 0.3 - 0.9 جـ- 0.18 - 0.22 د- (1-) - 4

هـ- (0.645 -) - (0.426 -) .

الحـل :

حيث أن الحدود الحقيقية للفترات يجب أن تتضمن رقم عشـري واحـد إضـافي عن القيـم المعطـاة وعليه فإن :

أ- نطرح 0.5 من بداية الفترة ونضيف 0.5 لنهايتها وبالتالي تكون الحدود الحقيقية لهذه الفترة كما يلي : 13.5 - 8.5 .

ب- نطرح 0.05 من بداية الفترة ونضيف 0.05 لنهايتها وبالتالي تكون الحدود الحقيقية لهذه الفترة كما يلي : 0.25 - 0.95 .

جـ- نطرح 0.005 من بداية الفترة ونضيف 0.005 لنهايتها وبالتالي تكون الحدود الحقيقية لهذه الفترة كما يلي : 0.175 - 0.225 .

د- نطرح (5 .0 -) من بداية الفترة ونضيف (0.5) لنهايتها وبالتالي تكون الحدود الحقيقية لهذه الفترة كما يلي : (1.5 -) - 4.5 .

هـ- نطرح (0.0005-) من بداية الفترة ونضيف (0.0005) لنهايتها وبالتالي تكون الحدود الحقيقية لهذه الفترة كما يلي : (0.6455 -) - (0.4265) .

مثال (12) : بفرض أن البيانات التالية تمثل الفترة الزمنية التي عمرتها عينة عشوائية من النضائد السائلة المنتجة من قبل أحد المصانع قبل أن تستهلك بالكامل :

1.3 2.1 0.3 2.5 4.3 1.8 1.4 2.0 1.9 1.7

2.8 3.7 3.1 2.3 1.5 2.6 3.5 5.9 2.0 1.2

2.1 0.2 1.1 2.8 1.3 2.3 3.4 1.8 2.6 3.9

0.7 2.9 3.5 0.9 2.1 2.4 0.4 3.9 6.3 2.5

1.2 5.3 1.7 2.7 1.8 0.4 4.6 3.2 1.6 2.4

والمطلوب :

أ- وضع البيانات في جدول توزيع تكراري ذي فترات منفصلة ثم أوجد الحدود الحقيقية للفترات والتكرار النسبي والمئوي .

ب - مثل البيانات بيانياً باستخدام المدرج التكراري والمضلع التكراري والمنحنى المتجمع الصاعد.

الحـل :

أ - لوضع هذه البيانات في جدول توزيع تكراري ذي فترات متصلة أو منفصلة نتبع الآتي :

1 - عدد الفترات المناسب : $k = 2.5\sqrt[4]{n} = 2.5\sqrt[4]{50} = 6.648 \approx 7.0$

وكما أشرنا سابقاً ليس بالضرورة اختيار نفس العدد الناتج عن هذه الطريقة ، ولكن إذا أردنا اختيار عدد مختلف عن هذا العدد يجب أن لا يكون مختلف بشكل كبير عن العدد الناتج عن هذه الطريقة وعليه سوف نختار عدد الفترات يساوى 8 . ألحظ أنه لو استخدمنا الطريقة الثانية فإن :

$$k = 1 + (3.322)\log_{10} 50 = 6.644 \approx 7.0$$

ومنها يتضح أن العدد متقارب في الطريقتين .

2 – تحديد المدى التي تنتشر فيه البيانات حيث المدى = 6.3 − 0.2 = 6.1 .

3 – طول الفترة = المدى ÷ عدد الفترات = 6.1 ÷ 8 = 0.7625 ≅ 0.8 تقريباً .

ألحظ أن : $8 \times 0.8 > 6.3$ ، إذن بعد إتمام هذه الخطوات يمكننا تكوين جدول التوزيع التكراري ذي الفترات المنفصلة وإيجاد الحدود الحقيقية لها كما يلي :

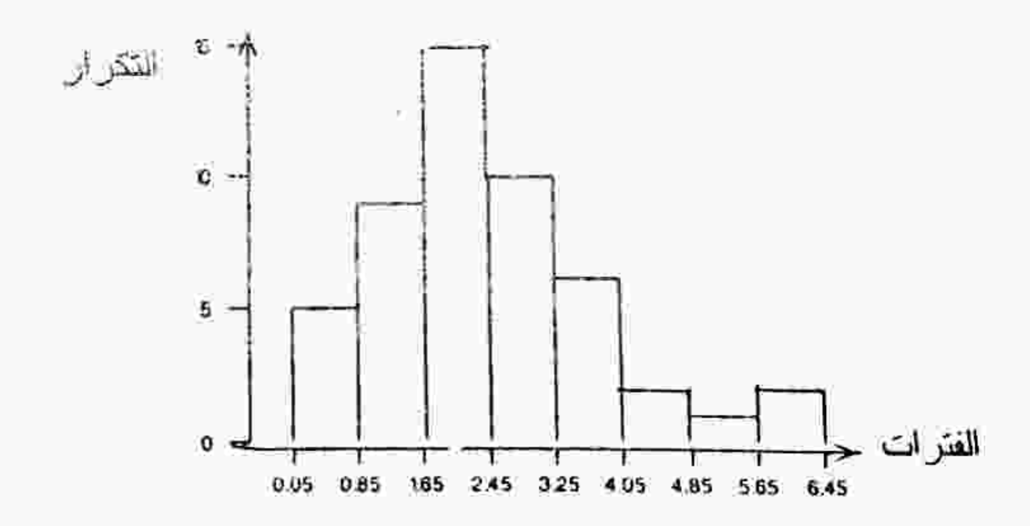
الفترات	التكرار f_i	الحدود الحقيقية	التكرار النسبي	التكرار المئوي
0.8 − 0.1	5	0.85 − 0.05	0.10	10
1.6 − 0.9	9	1.65 − 0.85	0.18	18
2.4 − 1.7	15	2.45 − 1.65	0.30	30
3.2 − 2.5	10	3.25 − 2.45	0.20	20
4.0 − 3.3	6	4.05 − 3.25	0.12	12
4.8 − 4.1	2	4.85 − 4.05	0.04	4
5.6 − 4.9	1	5.65 − 4.85	0.02	2
6.4 − 5.7	2	6.45 − 5.65	0.04	4
المجموع	50		1.0	100

ب – تمثيل البيانات بيانياً باستخدام المدرج التكراري، والمضلع التكراري والمنحى المتجمع الصاعد :

كما أشرنا سلفاً لتمثيل الجدول السابق بيانياً باستخدام المدرج التكراري نستخدم الحدود الحقيقية للفترات والتكرار المقابل لها بينما لتمثيله باستخدام المضلع التكراري نستخدم مراكز الفترات والتكرار المقابل لها ويمكن إيجاد مراكز الفترات إما من الحدود الظاهرية أو الحقيقية

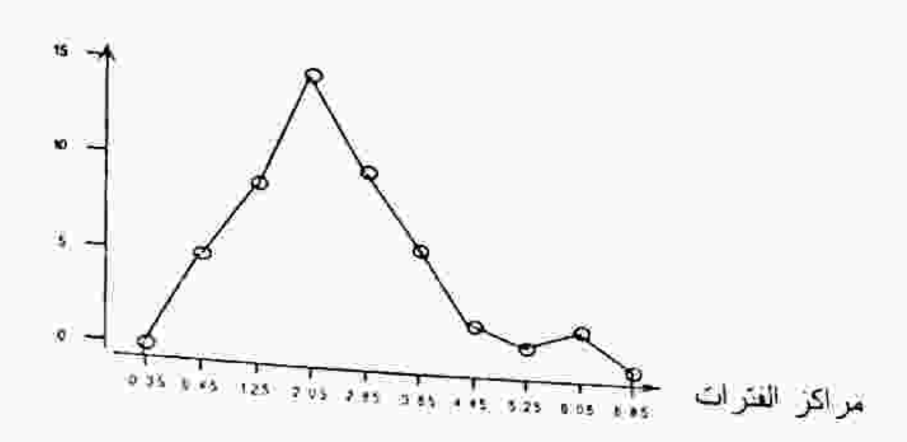
وذلك لأن الناتج واحد في كلتا الحالتين بينما لتمثيلـه باستخدام المنحـى المتجمـع الصـاعد نسـتخدم الحدود الحقيقية العليا للفترات والتكرار المتجمع الصاعد المقابل لها .

الفترات	مراكز الفترات	الحدود العليا للفترات	التكرار المتجمع الصاعد
0.8−0.1	0.45	أقل من 0.85	5
1.6−0.9	1.25	أقل من 1.65	14
2.4−1.7	2.05	أقل من 2.45	29
3.2−2.5	2.85	أقل من 3.25	39
4.0−3.3	3.65	أقل من 4.05	45
4.8−4.1	4.45	أقل من 4.85	47
5.6−4.9	5.25	أقل من 5.65	48
6.4−5.7	6.05	أقل من 6.45	50

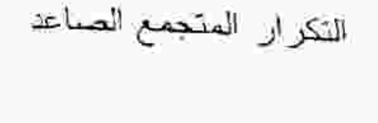


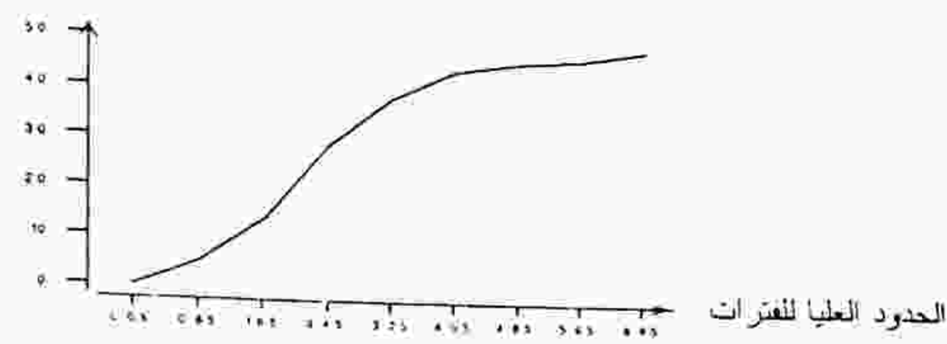
شكل (5) : المدرج التكراري





شكل (6) : المضلع التكراري





شكل (7) : المنحنى التكراري المتجمع الصاعد

العرض البياني في حالة الفترات غير متساوية الأطوال :

في بعض الأحيان قد يستحيل من الناحية العملية أو أن طبيعة البيانات لا تسمح بتكوين جدول توريع تكراري ذي فترات وبالتالي لتمثيل مثل هذا النوع من البيانات من الجداول بيانياً وبأي طريقة من الطرائق السابق ذكرها يجب تعديل التكرار بكل فترة وفقاً لطول تلك الفترة إلا في حالة المنحنى التكراري المتجمع الصاعد والهابط ، حيث التكرار المعدل لأي فترة يعرف كما يلي :

التكرار المعدل لأي فترة = التكرار بتلك الفترة ÷ طول الفترة نفسها .

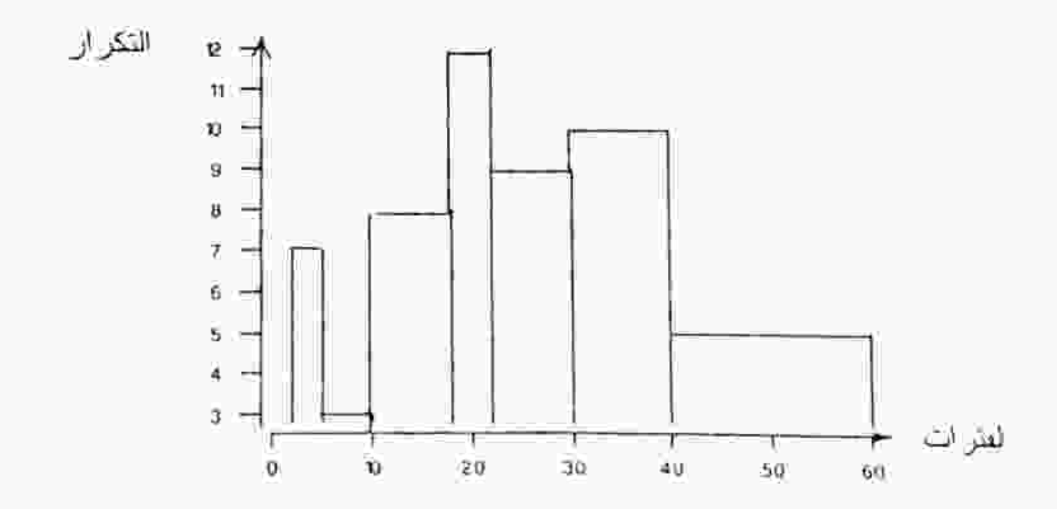
مثال (13): مثل البيانات الآتية بيانياً باستخدام المدرج التكراري

الفترات	2 -	5 -	10 -	18 -	22 -	30 -	40 - 60
التكرار f	21	15	64	98	72	100	100

الحـــل :

لتمثيل البيانات السابقة بيانياً باستخدام المـدرج التكراري يجب استخدام التكرار المعدل في التمثيل وذلك لأن أطوال الفترات غير متساوية .

الفترات	التكرار (f)	طول الفترة	التكرار المعدل
5 - 2	21	3	7
10 - 5	15	5	3
18 - 10	64	8	8
22 - 18	48	4	12
30 - 22	72	8	9
40 - 30	100	10	10
60 - 40	100	20	5



شكل (8) : المدرج التكراري

II -التمثيل البياني للبيانات النوعية

لوصف وتلخيص البيانات النوعية بيانياً حيث تكون الظاهرة قيد الدراسة مقسمة إلى صفات أو أنواع أو أزمنة يتم استخدام الأعمدة البيانية أو القطاعات الدائرية .

أ – الأعمدة البيانية Bar Chart

لتمثيل البيانات النوعية بيانياً باستخدام الأعمدة البيانية يتم تحديد المحور الأفقي لصفات أو أنواع أو أزمنة الظاهرة قيد الدراسة ، ويحدد المحور الرأسي للتكرارات المناظرة لها ثم يتم تمثيل كل نوع من أنواع الظاهرة قيد الدراسة بعمود بحيث تكون قواعد هذه الأعمدة متساوية ، على أن يكون عرض القاعدة للأعمدة يتناسب مع طبيعة البيانات وعادة ما يكون عرض القاعدة ما بين 0.5 سم و 1 سم ، ويجب ترك مسافة ما بين العمود والأخر هذه المسافة تساوى نصف قاعدة العمود تقريباً (يجب أن تكون المسافة بين الأعمدة متساوية) ، وعلى كل صفة يقام عموداً ارتفاعه يساوى التكرار لتلك الصفة وتتم المقارنة بين الأعمدة المختلفة من حيث كمية المعلومات التي يحتويها كل عمود من خلال ارتفاعه أي أنه عكس المدرج التكراري حيث تتم المقارنة هناك باستخدام المساحات المختلفة للمستطيلات .

والجدير بالملاحظة هنا أنه بالإمكان رسم أكثر من ظاهرة واحدة على نفس العمود وذلك إذا كان الغرض من الرسم مقارنة ظاهرتين أو أكثر بيانياً .

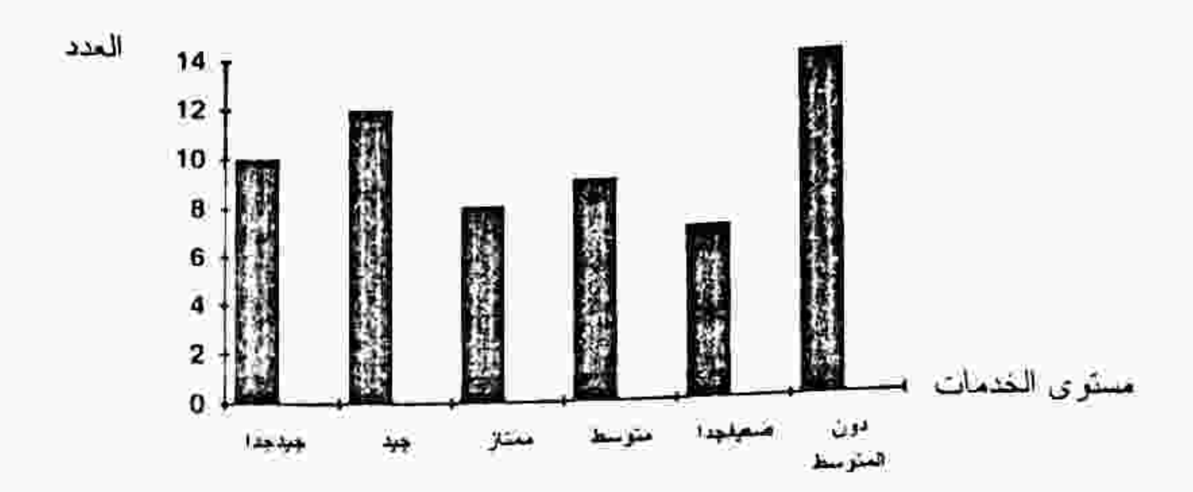
مثال (14) : بفرض أن البيانات الآتية تمثل آراء عينة من المرضى في استبيان قامت به أمانة الصحة حول مستوى الخدمات بهذا القطاع :

مستوى الخدمات	جيد جداً	جيد	ممتاز	متوسط	ضعيف جداً	دون المتوسط
عدد المرضى	10	12	8	9	7	14

والمطلوب تمثيل هذه البيانات بيانياً باستخدام الأعمدة البيانية .

الحــل :

يتم تمثيل هذه البيانات بيانياً باستخدام الأعمدة البيانية كما يلي :



شكل (9) : الاعمدة البيانية

ب - القطاعات الدائرية

تستخدم القطاعات الدائرية لوصف وتلخيص البينات النوعية ، حيث يتم رسم دائرة بنصف قطر معين ويتم تمثيل كل نوع من أنواع الظاهرة قيد الدراسة بجزء من تلك الدائرة يسمى قطاع ، ولرسم كل قطاع يتطلب الأمر معرفة زاويته ، وحيث أن مساحة القطاع بالدائرة يتناسب مع الزاوية المركزية ، (الزاوية المركزية للدائرة تساوى °360) وعليه فإن :

زاوية القطاع = °360 × التكرار النسبي للقطاع

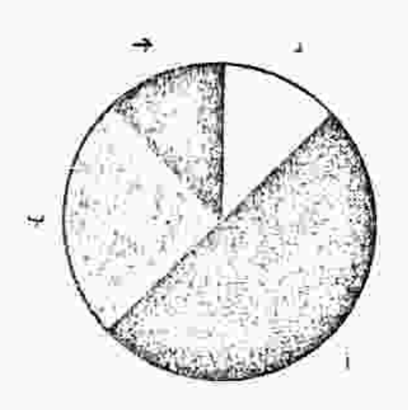
ويفضل هذا النوع من التمثيل البياني عن الأعمدة البيانية عندما تكون التكرارات كبيرة وذلك لأنه تصعب عملية تقسيم المحور الذي تقع عليه التكرارات ، بينما يفضل استخدام الأعمدة البيانية إذا كان الغرض من الرسم مقارنة ظاهرتين أو أكثر بيانياً .

مثال (15) : بفرض أن الجدول التالي يبين إنتاج مصنع المعمورة من 4 سلع مختلفة خلال فترة زمنية معينة والمطلوب تمثيل ذلك بيانياً باستخدام القطاعات الدائرية .

السلعة	الكمية
أ	400
ب	200
جـ	100
د	100
المجموع	800

الحـــل :

السلعة	الكمية	التكرار النسبي	زاوية القطاع
أ	400	0.5	360° × 0.5 = 180°
ب	200	0.25	360° × 0.25 = 90°
جـ	100	0.125	360° × 0.125 = 45°
د	100	0.125	369° × 0.125 = 45°
المجموع		1	360°



شكل (10) : القطاعات الدائرية

1 - 4 مقاييس النزعة المركزية Measures of Central Tendency

في البنود السابقة تم التطرق إلى وصف وتلخيص البيانات باستخدام الجداول التكرارية والرسومات البيانية وكل منها يعطى وصفاً عاماً وسريعاً للبيانات الإحصائية ، ولكن فوائدها الاستنتاجية محدودة جداً لذلك دعت الحاجة إلى وجود مقاييس عددية لوصف البيانات الإحصائية المتعلقة بالظاهرة قيد الدراسة .

وبالتمعن في القيم التي تأخذها الظواهر محل الدراسة نلحظ غالبية هذه القيم قريبة من بعضها البعض حيث نجد أن عدداً كبيراً من تلك القيم يميل إلى التجمع حول قيمة متوسطة ، أي قيمة غير منظورة تقع في وسط (مركز) البيانات وتعمل على جذب القيم إليها وكأن هناك نزعة عند البيانات للتجمع حول تلك القيمة ويقل عدد البيانات تدريجياً كلما ابتعدت البيانات عن تلك القيمة المتوسطة . لذلك سميت هذه الظاهرة الطبيعية بالنزعة المركزية (Central Tendency) .

وحيث أن التجمع حول هذه القيمة التي سيكون موقعها في الوسط ، فقد سميت بالمتوسط ، وذلك لأنها تتوسط هذا التجمع وتعبر عنه بصفة عامة . ومن خصائص المتوسط الجيد ما يلي :

أ - أن يكون معرف بشكل دقيق وقيمته تتوقف على الأعداد المستخرج منها .

ب - أن يأخذ في الحسبان جميع القيم بالمجموعة .　　جـ - أن يكون سهلاً وسريعاً في حسابه .

د - أن لا يتأثر كثيراً بالمتقلبات في قيم العينة .　　هـ - أن يخضع للعمليات الجبرية .

هناك عدة أنواع من المتوسطات منها المتوسط الحسابي والوسيط والمنوال والمتوسط الهندسي والمتوسط التوافقي واستعمال أي منها يعتمد على الهدف من الدراسة وطبيعة البيانات الإحصائية.

والجدير بالذكر أن هذه المقاييس لا تحل محل البيانات التفصيلية ولكنها تعطى فكرة واضحة عن الظاهرة قيد الدراسة ، وسوف نقوم بدراستها بعد التعرف على بعض الرموز والمصطلحات التي سوف نستخدمها في هذا البند ولاحقاً .

إذا كانت X ترمز لظاهرة (أو متغير) ما وتأخذ القيم $x_1 , x_2 , x_3 , \cdots , x_n$ فإن x_i ترمز لقيمة المفردة i حيث $i = 1 ، 2 ، 3 ، \ldots ، n$ وإن

أ - مجموع القيم نرمز له بالرمز $\sum_{i=1}^{n} x_i$ حيث $\sum_{i=1}^{n} x_i = x_1 + x_2 + x_3 + \cdots + x_n$.

ب - إذا ضربنا كل قيمة من القيم في مقدار ثابت وليكن a مثلاً فإن

$$\sum_{i=1}^{n} a\, x_i = a\, x_1 + a\, x_2 + \cdots + a\, x_n$$

$$= a\,(x_1 + x_2 + \cdots + x_n) = a \sum_{i=1}^{n} x_i$$

جـ - مجموع مربع كل قيمة يرمز له بالرمز : $\sum_{i=1}^{n} x_i^2 = x_1^2 + x_2^2 + x_3^2 + \cdots + x_n^2$.

د - مربع مجموع القيم يرمز له بالرمز : $(\sum_{i=1}^{n} x_i)^2 = (x_1 + x_2 + x_3 + \cdots + x_n)^2$.

هـ - إذا كانت جميع القيم متساوية وكل منها يساوى مقدار ثابت وليكن a مثلاً فإن

$$\sum_{i=1}^{n} a = a + a + a + \cdots + a = na$$

و - إذا كانت الظاهرة X تأخذ القيم $x_1 , x_2 , x_3 , \cdots , x_n$ والظاهرة Y تأخذ القيم $y_1 , y_2 , \cdots , y_n$ فإن

1- $$\sum_{i=1}^{n} x_i y_i = x_1 y_1 + x_2 y_2 + x_3 y_3 + \cdots + x_n y_n$$

2- $$\sum_{i=1}^{n} (x_i \pm y_i) = \sum_{i=1}^{n} x_i \pm \sum_{i=1}^{n} y_i$$
$$= (x_1 + x_2 + x_3 + \ldots + x_n) \pm (y_1 + y_2 + y_3 + \ldots + y_n)$$

3- $$\sum_{i=1}^{n} (a x_i \pm b y_i) = a \sum_{i=1}^{n} x_i \pm b \sum_{i=1}^{n} y_i$$

حيث a و b ثوابت .

4- $$\sum_{i=1}^{n} \frac{x_i}{y_i} = \frac{x_1}{y_1} + \frac{x_2}{y_2} + \ldots + \frac{x_n}{y_n}$$.

5- $$\frac{\sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} y_i} = \frac{x_1 + x_2 + \ldots + x_n}{y_1 + y_2 + \ldots + y_n}$$.

مثال (16): إذا كانت المفردات 2 ، 8 ، 4 ، 5 تمثل قيم ظاهرة ما فإنه يمكن تمثيلها بدلالة الرموز كالآتي :

$x_1 = 2$, $x_2 = 8$, $x_3 = 4$, $x_4 = 5$ وإن

1) $$\sum_{i=1}^{4} x_i = x_1 + x_2 + x_3 + x_4 = 2+8+4+5 = 19$$

2) $$\sum_{i=1}^{4} x_i^2 = x_1^2 + x_2^2 + x_3^2 + x_4^2 = (2)^2 + (8)^2 + (4)^2 + (5)^2$$
$$= 4 + 64 + 16 + 25 = 109$$

3) $$(\sum_{i=1}^{4} x_i)^2 = (x_1 + x_2 + x_3 + x_4)^2 = (2+8+4+5)^2 = (19)^2 = 361$$

4) $$\sum_{i=1}^{4} 6 x_i = 6 \sum_{i=1}^{4} x_i = 6(19) = 114$$

5) إذا كانت المفردات 3، 5 ، 6 ، 7

ولتكن Y فإن

أ -

$$\sum_{i=1}^{4} x_i\, y_i = (2)(3)+(8)(5)+(4)(6)+(5)(7) = 6+40+24+35 = 105$$

ب -

$$\sum_{i=1}^{4} (x_i + y_i) = (2+3)+(8+5)+(4+6)+(5+7) = 5+13+10+12 = 40$$

أو

$$\sum_{i=1}^{4} (x_i + y_i) = \sum_{i=1}^{4} x_i + \sum_{i=1}^{4} y_i = 19+21 = 40$$

جـ-

$$\frac{\sum_{i=1}^{4} x_i}{\sum_{i=1}^{4} y_i} = \frac{19}{21} = 0.9048$$

د-

$$\sum_{i=1}^{4} \frac{x_i}{y_i} = \frac{2}{3}+\frac{8}{5}+\frac{4}{6}+\frac{5}{7} = 3.648$$

1 - 4 - 1 المتوسط الحسابي Arithmetic Mean

يستخدم هذا المقياس إذا كان الهدف من الدراسة الحصول على مقياس مستقر لا يختلف كثيراً من عينة لأخرى أو عندما يكون الهدف القيام بتحليل إحصائي مثل تقدير معلمات المجتمع أو اختبار الفرضيات الإحصائية كما سنرى في ما بعد .

أ - طريقة حساب المتوسط الحسابي من البيانات الأولية

(1) الطريقة المباشرة Direct Method

إذا كانت $x_1, x_2, x_3, \cdots, x_n$ تمثل قيم لظاهرة ما فإن المتوسط الحسابي لهذه القيم هو مجموع هذه القيم مقسوماً على عددها ويرمز له بالرمز $\bar{x}$ أي أن

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} \qquad (1)$$

حيث n تمثل عدد القيم .

مثال(17) : بفرض أن القيم 30 ، 25 ، 20 ، 26 ، 28 ، 24 ، 28 ، 27 تمثل المبالغ التي أنفقها ثمانية زبائن عند تناولهم لوجبة غذاء بأحد الفنادق والمطلوب حساب متوسط الإنفاق .

الحـل :

$$\bar{x} = \frac{\sum_{i=1}^{8} x_i}{8} = \frac{30+25+20+26+28+24+28+27}{8} = \frac{208}{8} = 26$$

أي أن متوسط الأنفاق على تناول وجبة غذاء بهذا الفندق هو 26 دينار .

(2) طريقة الانحرافات البسيطة(طريقة المتوسط الفرضي) Normal Deviation Method

إذا كانت $x_1, x_2, x_3, \dots, x_n$ تمثل مفردات لظاهرة معينة وكان العدد A يمثل المتوسط الفرضي لهذه البيانات (يفضل أن تكون قيمة A مساوية لإحدى المفردة التي بالظاهرة وأن تكون هذه المفردات واقعة في المنتصف من حيث قيمتها) وكانت $d_i = x_i - A$ أي أن d_i تمثل انحرافات المفردات عن وسطها الفرضي فإن المتوسط الحسابي في هذه الحالة معرف كما يلي:-

$$\bar{x} = A + \frac{\sum_{i=1}^{n} d_i}{n} \qquad (2)$$

$$\Rightarrow \qquad \bar{x} = A + \bar{d}$$

وذلك لأن: $d_i = x_i - A$ وعليه فإن

$$\sum_{i=1}^{n} d_i = \sum_{i=1}^{n} (x_i - A) = \sum_{i=1}^{n} x_i - \sum_{i=1}^{n} A = \sum_{i=1}^{n} x_i - nA$$

$$\Rightarrow \quad \frac{\sum_{i=1}^{n} d_i}{n} = \frac{\sum_{i=1}^{n} x}{n} - \frac{nA}{n} = \bar{x} - A \quad \Rightarrow \quad \bar{x} = \frac{\sum_{i=1}^{n} d_i}{n} + A = \bar{d} + A$$

تستخدم هذه الطريقة إذا كانت البيانات كثيرة وقيمها كبيرة بحيث يصعب إيجاد متوسطها بالطريقة المباشرة .

مثال (18): أوجد المتوسط الحسابي باستخدام الوسط الفرضي للبيانات الآتية :

10 ، 9 ، 12 ، 11 ، 8 ، 13 ، 14 ، 10 ، 15

الحـل :

باختيار الوسط الفرضي : $A = 11$ وعليه فإن

$$\sum_{i=1}^{9} d_i = \sum_{i=1}^{9} (x_i - A)$$

$$= (10-11) + (9-11) + (12-11) + (11-11)$$

$$+ (8-11) + (13-11) + (14-11) + (10-11) + (15-11)$$

$$= -1 - 2 + 1 + 0 - 3 + 2 + 3 - 1 + 4 = 3$$

$$\overline{x} = A + \frac{\sum_{i=1}^{n} d_i}{n} \quad \Rightarrow \quad \overline{x} = 11 + \frac{3}{9} = 11.333$$

(3) طريقة الانحرافات المختصرة Short - Cut Deviation Method

يتم حساب المتوسط الحسابي في هذه الحالة كما يلي :

$$\overline{x} = A + c\overline{u} \quad , \quad u_i = \frac{d_i}{c} \qquad (3)$$

حيث A وكما سبق ترمز للمتوسط الفرضي و $d_i = x_i - A$ و c يطلق عليه تسمية ثابت القسمة ، وهو أكبر قيمة عددية تقبل كل القيم القسمة عليه وبدون باقي (كسر) كلما أمكن ذلك من أجل تبسيط العمليات الحسابية .

مثال (19) : إذا كانت البيانات التالية : 25 ، 20 ، 10 ، 15 ، 30 تمثل درجات 5 طلاب في امتحان لمادة الإحصاء .أوجد متوسط درجات هذا الامتحان .

الحـل :

باختيار الوسط الفرضي : $A = 15$ و $c = 5$ نجد أن

$d_i = x_i - A$	30 - 15 = 15	15 - 15 = 0	10 - 15 = -5	20 - 15 = 6	25 - 15 = 10
$u_i = \frac{d_i}{5}$	$\frac{15}{5} = 3$	$\frac{0}{5} = 0$	$\frac{-5}{5} = -1$	$\frac{5}{5} = 1$	$\frac{10}{5} = 2$

وبالتالي فإن $\bar{u} = \frac{\sum_{i=1}^{5} u_i}{5} = \frac{5}{5} = 1$ و $\bar{x} = A + c\bar{u} = 15 + (5)(1) = 20$.

ب – طريقة حساب المتوسط الحسابي من الجداول التكرارية

(1) الطريقة المباشرة Direct Method

تفيد هذه الطريقة في حالة ما يكون جدول التوزيع التكراري بفترات متساوية في الطول أو غير متساوية . فإذا كانت $x_1, x_2, x_3, \dots, x_n$ تمثل قيم الظاهرة x و $f_1, f_2, f_3, \dots, f_k$ تمثل التكرارات المناظرة لها فإن المتوسط الحسابي معرف كما يلي :-

$$\bar{x} = \frac{\sum_{i=1}^{k} f_i x_i}{n} \qquad (4)$$

حيث :

x_i = المفردة i (أو مركز الفترة i) ، f_i = تكرار المفردة i (أو تكرار الفترة i)

k = عدد المفردات المختلفة عن بعضها البعض (أو عدد الفترات) .

$n = \sum_{i=1}^{k} f_i$ = مجموع التكرارات .

مثال (20): يوجد بأحد المستشفيات موظفاً يتقاضى راتباً شهرياً قدره 200 دينار وخمسة موظفين يتقاضى كل منهم 180 دينار شهرياً و10 موظفين يتقاضى كل منهم 140 دينار شهرياً والمطلوب إيجاد متوسط الرواتب الشهرية في هذا المستشفى .

الحـل :

بفرض أن x ترمز للرواتب المختلفة و f ترمز لعدد الموظفين فإنه بالإمكان تلخيص العمليات الحسابية في الجدول التالي :

المرتب (x_i)	عدد الموظفين (التكرار f_i)	$f_i x_i$
200	1	200
180	5	900
140	10	1400
المجموع	16	2900

$$\overline{x} = \frac{\sum_{i=1}^{k} f_i x_i}{n} = \frac{2500}{16} = 156.25$$

ألحظ أنه إذا اختلفت الأهمية النسبية للبيانات حيث يكون لكل مفردة وزناً يختلف عن بقية المفردات ، يتم استخدام متوسط حسابي يطلق عليه تسمية المتوسط الحسابي المرجح (الموزون) (weighted mean) ، فإذا كان لدينا بيانات ولتكن $x_1, x_2, x_3, \cdots, x_n$ وأوزانها على الترتيب هي $w_1, w_2, w_3, \cdots, w_n$ فإن المتوسط المرجح لهذه البيانات سنرمز له بالرمز $\overline{w}$ ، ومعرف كما يلي :

$$\overline{w} = \frac{\sum_{i=1}^{n} w_i x_i}{\sum_{i=1}^{n} w_i} \qquad (5)$$

مثال (21) :الجدول التالي يمثل المقررات الدراسية وعدد الوحدات لكل منها والدرجة المتحصل عليها لطالب ما في إحدى الفصول الدراسية بقسم الإحصاء بكلية العلوم بجامعة الفاتح والمطلوب حساب المتوسط العام لدرجات هذا الطالب خلال هذا الفصل :

الدرجة	عدد الوحدات	المقرر
50	3	ST 403
62	4	ST 402
80	4	ST 410
55	4	ST 405
70	2	ST 492

الحل :

من تعريف المتوسط الحسابي المرجح نجد أن

$$\overline{w} = \frac{3\times 50 + 4\times 62 + 4\times 80 + 4\times 55 + 2\times 70}{15} = \frac{1078}{15} = 71.867$$

مثال (22) : أوجد متوسط الإنفاق مستخدماً بيانات المثال رقم (7) .

الحل :

لحساب المتوسط الحسابي في هذه الحالة يتطلب أولا أيجاد مراكز الفترات ثم نوجد حاصل ضرب مراكز الفترات في التكرار المقابل لها كما يتضح في الجدول التالي :

الفترات	التكرار (f_i)	مراكز الفترات (x_i)	$f_i x_i$
56 - 50	1	53	53
62 - 56	5	59	295
68 - 62	12	65	780
74 - 68	15	71	1065
80 - 74	22	77	1694
86 - 80	11	83	913
92 - 86	6	89	534
98 - 92	3	95	285
المجموع	$\sum_{i=1}^{8} f_i = 75$		$\sum_{i=1}^{8} f_i x_i = 5619$

إذن متوسط الأنفاق سيكون كالأتي :-

$$\bar{x} = \frac{\sum_{i=1}^{8} f_i x_i}{\sum_{i=1}^{8} f_i} = \frac{5619}{75} = 74.92$$

أي أن متوسط الانفاق يساوى 74.92 دينار .

(2) طريقة الانحرافات البسيطة

في هذه الحالة يتم حساب المتوسط الحسابي كما يلي :

$$\bar{x} = A + \frac{\sum_{i=1}^{k} f_i d_i}{n} \qquad (6)$$

حيث A ترمز للوسط الفرضي و $d_i = x_i - A$ ترمز لانحرافات مراكز الفترات (x_i) عن الوسط الفرضي .

ألحظ أن الوسط الفرضي هنا يفضل أن تكون قيمته المفردة ذات الأكبر تكرار في حالة البيانات المكررة ، ومركز الفترة التي تقع في وسط الفترات وان يكون لتلك الفترة أكبر تكرار إن أمكن ذلك .

مثال (23): أوجد المتوسط الحسابي باستخدام الوسط الفرضي لبيانات المثال (7) .

الحـــل :

باختيار الوسط الفرضي مساوياً لمركز الفترة التي يقابلها أكبر تكرار نجد أن

الفترات	التكرار (f_i)	مراكز الفترات x_i	$d_i = x_i - A$	$f_i d_i$
-50	1	53	-24	- 24
-56	5	59	- 18	- 90
-62	12	65	- 12	- 144
-68	15	71	- 6	- 90
-74	22	77	0	0
-80	11	83	6	66
-86	6	89	12	72
98-92	3	95	18	54
المجموع	75			$\sum_{i=1}^{n} f_i d_i = -156$

وعليه فإن المتوسط الحسابي يكون كالآتي :

$$\bar{x} = 77 + \left(\frac{-156}{75}\right) = 77 - 2.08 = 74.92$$

(3) طريقة الانحرافات المختصرة

في هذه الحالة يتم حساب المتوسط الحسابي كالآتي :

$$\overline{x} = A + c\left(\frac{\sum_{i=1}^{k} f_i d_i}{n} \right) = A + c\overline{u} \qquad (7)$$

حيث $u_i = \frac{d_i}{c}$ ، و c ثابت القسمة ويفضل أن تكون قيمته مساوية لأكبر قيمة تقبل قيم الظاهرة القسمة عليها وبدون باقي (كسر) إن أمكن ذلك في البيانات المكررة ، كما يفضل أن تكون قيمته مساوية لطول الفترة في حالة الجداول التكرارية ذات الأطوال المتساوية .

مثال (24) : من نفس بيانات المثال السابق أوجد المتوسط الحسابي باستخدام طريقة الانحرافات المختصرة .

الحل :

حيث أن طول الفترة يساوى 6 وعليه فإن c = 6 وباختيار A = 77 وذلك لأنها تقابل أكبر تكرار نجد أن

الفترات	التكرار (f_i)	مراكز الفترات x_i	$d_i = x_i - A$	$u_i = \frac{d_i}{c_i}$	$f_i u_i$
-50	1	53	- 24	- 4	- 24
-56	5	59	- 18	- 3	- 90
-62	12	65	- 12	- 2	- 144
-68	15	71	- 6	- 1	- 90
-74	22	77	0	0	0
-80	11	83	6	1	66
-86	6	89	12	2	72
98-92	3	95	18	3	54
المجموع	75				- 26

وبالتالي يكون

$$\overline{x} = A + c\overline{u} = 77 + 6\left(\frac{-26}{75}\right) = 77 - 2.08 = 74.92$$

بعض خواص المتوسط الحسابي :

1 - مجموع انحرافات القيم عن متوسطها الحسابي يساوي صفر . ولإثبات ذلك ألحظ أن

$$\sum_{i=1}^{n}(x_i - \bar{x}) = \sum_{i=1}^{n} x_i - \sum_{i=1}^{n} \bar{x} = n\left(\frac{\sum_{i=1}^{n} x_i}{n}\right) - n\bar{x} = n\bar{x} - n\bar{x} = 0$$

2 - مجموع مربعات انحرافات القيم عن متوسطها الحسابي أصغر من مجموع مربعات انحرافات القيم عن أي قيمة أخرى .

فإذا افترضنا أن a أي مقدار ثابت يختلف عن المتوسط الحسابي ، فإن ما تتضمنه هذه الخاصية هو

$$\sum_{i=1}^{n}(x_i - \bar{x})^2 < \sum_{i=1}^{n}(x_i - a)^2$$

ولإثبات ذلك سوف نوجد المشتقة الأولى للطرف الأيمن ومساواة هذه المشتقة بالصفر أي أن

$$\frac{d}{da}\sum_{i=1}^{n}(x_i - a)^2 = -2\sum_{i=1}^{n}(x_i - a) = 0$$

$$\Rightarrow \quad \sum_{i=1}^{n} x_i - na = 0 \quad \Rightarrow \quad a = \frac{\sum_{i=1}^{n} x_i}{n} = \bar{x}$$

ألحظ أن $\sum_{i=1}^{n}(x_i - a) = 0$ لا يمكن أن يكون صحيحاً إلا إذا كانت $a = \bar{x}$ ؛ وذلك من الخاصية (1) للمتوسط الحسابي .

3 - قابل للعمليات الجبرية .

4 - لا يمكن حسابه بيانياً .

5 - لا يمكن حسابه من الجداول التكرارية المفتوحة ، وذلك لأنه يعتمد في حسابه على مراكز الفترات وإذا كانت الفترة مفتوحة لا يمكن تحديد مركزها ، ويقال بأن جدول التوزيع التكراري مفتوح إذا كانت بداية الفترة الأولى أو نهاية الفترة الأخيرة بالجدول غير محددة كما ما هو موضح في (أ) و (ب) على سبيل المثال .

أ - جدول توزيع تكراري مفتوح أو غير محدد من البداية :

الفترات	أقل من 5	5 -	10 -	15 -	15 - 20
التكرار f_i	4	7	11	8	6

ب - جدول توزيع تكراري مفتوح أو غير محدد من النهاية :

الفترات	3 –	7 –	11 –	15 –	19 –	23 فأكثر
التكرار f_i	20	11	19	8	12	15

6 - يتأثر بالقيم المتطرفة (الشاذة) ، ويقال بأن مفردة ما متطرفة إذا كانت كبيرة جداً أو صغيرة جداً مقارنة ببقية المفردات ، فمثلاً إذا كانت 7 ، 9 ، 6 ، 30 تمثل مفردات لظاهرة ما وكانت 45 ، 42 ، 48 ، 11 ، 39 تمثل مفردات لظاهرة أخرى ، فإنه يقال بأن المفردتان 30 و 11 مفردتين متطرفتين وذلك لأن المفردة 30 كبيرة مقارنة بمجموعتها بينما المفردة 11 صغيرة مقارنة بمجموعتها ، وبالتالي فإن قيمة المتوسط الحسابي أقل مما يجب إذا كانت المفردة صغيرة مقارنة ببقية القيم أو أن تكون قيمته أكبر مما يجب إذا كانت المفردة كبيرة مقارنة ببقية القيم . فمثلاً من البيانات الأولى نجد أن $\overline{x} = 13$ ومن البيانات الثانية نجد أن $\overline{x} = 37$.

7- عند حساب المتوسط الحسابي باستخدام المتوسط الفرضي يجب استخدام التكرار المعدل إذا كانت الفترات ذات أطوال غير متساوية .

8 - إذا كانت الظاهرة X تأخذ القيم $x_1, x_2, x_3, \cdots, x_m$ و $\overline{x}$ تمثل المتوسط الحسابي لها وكانت الظاهرة Y تأخذ القيم $y_1, y_2, \cdots, y_n$ و $\overline{y}$ تمثل المتوسط الحسابي لها فإن المتوسط الحسابي للظاهرتين معاً يمكن أن نرمز له بالرمز $\overline{\overline{x}}$ ومعرف كما يلي :

$$\overline{\overline{x}} = \frac{m\overline{x} + n\overline{y}}{m+n}$$

9 - المتوسط الحسابي أقل تفاوتاً من عينة لأخرى مقارنة ببقية مقاييس النزعة المركزية الأخرى (كما سنرى فيما بعد) .

10 - يُعد المتوسط الحسابي أهم مقاييس النزعة المركزية وأكثرها استخداماً .

1 - 4 - 2 المتوسط الهندسي Geometric Mean

المتوسط الهندسي لمجموعة من القيم والتي عددها n هو عبارة عن الجذر النوني (n^{th} root) لحاصل ضربها.

أ - حساب المتوسط الهندسي في حالة البيانات الأولية

إذا كانت البيانات $x_1, x_2, x_3, \ldots, x_n$ تمثل قيم لظاهرة ما فإن المتوسط الهندسي لهذه البيانات سنرمز له بالرمز G ويتم حسابه كما يلي :

$$G = \sqrt[n]{x_1 . x_2 . x_3 ... x_n} \qquad = \left[x_1 . x_2 . x_3 ... x_n \right]^{\frac{1}{n}} \qquad (8)$$

ويمكن استخدام اللوغاريتمات لحساب المتوسط الهندسي وذلك على النحو التالي :

$$\log_e G = \frac{1}{n} \sum_{i=1}^{n} \log x_i$$

أي أن لوغاريتم المتوسط الهندسي للأساس الطبيعي يساوى المتوسط الحسابي للوغاريتم القيم ، وبالبحث عن العدد المقابل للوغاريتم يتم الحصول على قيمة المتوسط الهندسي ، أي أن

$$G = e^{\frac{1}{n} \sum_{i=1}^{n} \log x_i} \qquad (9)$$

حيث $n = \sum_{i=1}^{k} f_i$.

مثال (25) : أوجد المتوسط الهندسي للبيانات التالية : 2 ، 4 ، 8 .

الحــل :

من التعريف أعلاه للمتوسط الهندسي نجد أن

$$G = \sqrt[n]{x_1 . x_2 ... x_n} \quad = \sqrt[3]{2 \times 4 \times 8} \quad = \sqrt[3]{64} \quad = (64)^{\frac{1}{3}} = 4$$

كما يمكن الحصول على النتيجة نفسها باستخدام اللوغاريتمات وذلك كما يلي :

$$\log_e G = \frac{1}{n} \sum_{i=1}^{n} \log x_i$$

$$= \frac{1}{3} \left[\log_e 2 + \log_e 4 + \log_e 8 \right]$$

$$= \frac{1}{3} \left[0.3010 + 0.6920 + 0.9031 \right] = \frac{1}{3} (18062) = 0.60207$$

$$\Rightarrow G = e^{\frac{1}{n} \sum_{i=1}^{n} \log x_i} = e^{0.60207} = 4$$

ألحظ بهذا المثال أن التناسب ما بين عددين متتاليين متساوي .

مثال (26) : بفرض أن إنتاج مصنع القره بوللى للدائن أرتفع بنسبة 2 ٪ خلال الفترة 95 - 1996 م ، كما أرتفع بنسبة 5 ٪ خلال الفترة 96 - 1997 م ، وأيضاً أرتفع بنسبة 10 ٪ في الفترة 97 - 1998 م . والمطلوب إيجاد معدل الزيادة في الإنتاج خلال الفترة 1995 إلى 1998 م .

الحـــل :

$$G = \sqrt[n]{x_1 \cdot x_2 \cdots x_n} = \sqrt[3]{\frac{102}{100} \times \frac{105}{100} \times \frac{110}{100}}$$

$$= \sqrt[3]{1.02 \times 1.05 \times 1.10} = \sqrt[3]{1.1781} = 1.056$$

وعليه فإن معدل الزيادة في الانتاج خلال الفترة من 1995 إلى 1998 يساوي 0.056 .

ب - حساب المتوسط الهندسي في حالة البيانات المكررة والمبوبة

إذا كانت $x_1, x_2, x_3, \ldots, x_k$ تمثل المفردات المختلفة للظاهرة x (أو مراكز الفترات) و $f_1, f_2, f_3, \ldots, f_k$ تمثل التكرارات المناظرة لها فإن المتوسط الهندسي لهذه البيانات يُعرف كما يلي :

$$G = \sqrt[n]{x_1^{f_1} \cdot x_2^{f_2} \cdot x_3^{f_3} \ldots x_k^{f_k}}$$

$$= \left[x_1^{f_1} \cdot x_2^{f_2} \cdot x_3^{f_3} \ldots \ldots x_k^{f_k} \right]^{\frac{1}{n}} \qquad (10)$$

وباستخدام اللوغاريتمات نجد أن :

$$\log_e G = \frac{1}{n} \sum_{i=1}^{k} f_i \log x_i \quad \Rightarrow \quad G = e^{\frac{1}{n} \sum_{i=1}^{k} f_i \log x_i} \qquad (11)$$

مثال (27) : أوجد المتوسط الهندسي لجدول التوزيع التكراري الآتي :

الفترات	-2	-4	-6	-8	$12-10$	المجموع
التكرار f_i	6	3	5	4	2	$\sum_{i=1}^{k} f_i = 20$

الحـل :

الفترات	التكـرار (f_i)	مراكز الفترات x_i	$(x_i)^{f_i}$
-2	6	3	729
-4	3	5	125
-6	5	7	16807
-8	4	9	6561
12-10	2	11	121

$$G = [729 \times 125 \times 16807 \times 6561 \times 121]^{\frac{1}{20}} = 5.7$$

ويمكن استخدام اللوغاريتمات في حساب المتوسط الهندسي وذلك كالآتي :

الفترات	التكرار (f_i)	مراكز الفترات (x_i)	$\text{Log } x_i$	$f_i \log x_i$
-2	6	3	0.4771	2.8626
-4	3	5	0.6990	2.097
-6	5	7	0.8451	4.2255
-8	4	9	0.9542	3.8168
12-10	2	11	1.0414	2.0828
المجموع	20			15.0847

وعليه فإن

$$\log G = \frac{\sum_{i=1}^{5} f_i \log x_i}{\sum_{i=1}^{5} f_i} = \frac{15.0847}{20} = 0.7542 \Rightarrow G = e^{0.7542} = 5.6786$$

بعض خواص المتوسط الهندسي :

1 - يدخل في حسابه جميع القيم .

2 - لا يتأثر كثيراً بالقيم المتطرفة .

3 - ليس له معنى إذا كانت إحدى القيم سالبة أو تساوي صفر .

4 - لا يمكن حسابه من الجداول التكرارية المفتوحة وذلك لأنه يعتمد في حسابه على مراكز الفترات .

5 - قابل للعمليات الجبرية .

6 - المتوسط الهندسي أصغر من المتوسط الحسابي لأي مجموعة من البيانات الموجبة وغير المتساوية .

7 - يستخدم في حساب الأرقام القياسية لمناسيب الأسعار، وعند تقدير عدد السكان بين سنتي التعداد .

1 - 4 - 3 المتوسط التوافقي Harmonic Mean

المتوسط التوافقي لمجموعة من البيانات هو عبارة عن مقلوب المتوسط الحسابي لمقلوب القيم . ويفضل استخدام هذا المتوسط في حساب معدل السرعة إذ إنها تعطى في العادة بدلالة وحدة الزمن وكذلك في حساب متوسط الأسعار متى أعطيت على أساس عدد الوحدات بالنسبة لوحدة النقود .

أ- طريقة حساب المتوسط التوافقي من البيانات الأولية

إذا كانت $x_1, x_2, x_3, ..., x_n$ تمثل قيم لظاهرة ما فإن المتوسط التوافقي لهذه القيم والذي سنرمز له بالرمز H سيكون كما يلي :-

$$H = \frac{n}{\sum_{i=1}^{n} \frac{1}{x_i}} \qquad (12)$$

مثال (28) : أوجد المتوسط التوافقي للقيم : 8 ، 4 ، 10 ، 2 ، 6 .

الحـــل :

من التعريف نجد أن

$$H = \frac{5}{\frac{1}{8} + \frac{1}{4} + \frac{1}{10} + \frac{1}{2} + \frac{1}{6}} = \frac{5}{1.1417} = 4.3794$$

ب- طريقة حساب المتوسط التوافقي من الجداول التكرارية

إذا كانت $x_1, x_2, x_3, ..., x_k$ تمثل المفردات المختلفة للظاهرة x (أو مراكز الفترات) و $f_1, f_2, f_3, ..., f_k$ تمثل التكرارات المناظرة لها فإن المتوسط التوافقي لهذه البيانات يُعرف كما يلي :

$$H = \frac{\sum_{i=1}^{k} f_i}{\sum_{i=1}^{k} \left(\frac{f_i}{x_i} \right)} \qquad (13)$$

مثال (29) : أوجد المتوسط التوافقي من الجدول الآتي :

الفترات	-3	-7	-11	-15	-19	27-23
التكرار	8	12	10	9	6	5

الحل :

الفترات	التكرار f_i	مراكز الفترات (x_i)	$\frac{f_i}{x_i}$
-3	8	5	1.6
-7	12	9	1.3333
-11	10	13	0.7682
-15	9	17	05294
-19	6	21	0.2857
27-23	5	25	0.2
المجموع	50		$\sum_{i=1}^{6} \left(\frac{f_i}{x_i} \right) = 4.7176$

إذن

$$H = \frac{\sum_{i=1}^{6} f_i}{\sum_{i=1}^{6}\left(\frac{f_i}{x_i}\right)} = \frac{50}{4.7176} = 10.5986$$

بعض خواص المتوسط التوافقي :

1 - يتأثر بالقيم الشاذة .

2 - لا يمكن حسابه من الجداول التكرارية المفتوحة .

3 - قابل للعمليات الجبرية .

4 - يستخدم في وصف تغيرات الظواهر النسبية وخاصة التي تميز معطياتها عكس تمييز المتوسط .

5 - يكون $H \leq G \leq \bar{x}$ ويحدث التساوي في حالة ما تكون جميع القيم متساوية .

1 - 4 - 4 الوسيط Median

الوسيط لمجموعة من البيانات هو القيمة التي تقسم مجموعة البيانات المرتبة ترتيباً تصاعدياً أو تنازلياً إلى قسمين متساويين بحيث يكون عدد البيانات الأصغر منها يساوى عدد البيانات الأكبر منها .

ويستخدم الوسيط كمقياس للنزعة المركزية إذا أردنا معرفة البيانات من حيث الموقع أو إذا أردنا وصفاً عاماً سريعاً للظاهرة قيد الدراسة ، ويستخدم في الحالات التي يتعذر فيها استخدام المتوسط الحسابي كما في حالة الجداول التكرارية المفتوحة ، أيضاً في الحالات آلتي يكون فيها المتوسط الحسابي مضللاً كما هو الحال عند وجود قيم شاذة في البيانات .

أ - طريقة حساب الوسيط من البيانات الأولية

لحساب الوسيط من البيانات الأولية يجب أولاً ترتيب البيانات ترتيباً تصاعدياً أو تنازلياً وهناك حالتين :-

1 - إذا كان عدد القيم (n) فردى فإن رتبة الوسيط = $\frac{1+n}{2}$. وبالتالي تكون قيمة الوسيط ($\tilde{X}$) هي المفردة آلتي ترتيبها يقابل رتبة الوسيط .

2 - إذا كان عدد القيم (n) زوجي فإن رتبة الوسيط = ($\frac{n}{2}$ و $\frac{n}{2}+1$) وبالتالي فإن قيمة الوسيط ($\tilde{x}$) تساوى المتوسط الحسابي للمفردتين المناظرتين لرتبة الوسيط .

مثال (30): أوجد الوسيط مستخدماً بيانات المثال (17) .

الحل :

لحساب الوسيط يجب أولاً ترتيب القيم ترتيباً تصاعدياً أو تنازلياً وذلك على النحو التالي :

رتب القيم	8	7	6	5	4	3	2	1
القيم	30	28	28	27	26	25	24	20

حيث أن عدد القيم (n) = 8 زوجي وبالتالي فإن

رتبة الوسيط = ($\frac{n}{2}$ و $1+\frac{n}{2}$) = ($4=\frac{8}{2}$ و $5=1+4$) وعليه فإن

$$\tilde{x} = \frac{26+27}{2} = \frac{53}{2} = 26.5$$

ألحظ أنه : إذا كان عدد القيم هنا سبعة فقط أي لو استبعدنا المفردة 30 مثلاً فإن الوسيط ($\tilde{x}$) يساوى 26 (لماذا ؟) .

ب - طريقة حساب الوسيط من الجداول التكرارية

(1) الطريقة الحسابية:

لحساب الوسيط من جدول توزيع تكراري نتبع الآتي :-

1 - حساب التكرار المتجمع الصاعد .

2 - حساب رتبة الوسيط حيث رتبة الوسيط = $\frac{\sum_{i=1}^{k} f_i}{2}$ ، بغض النظر فيما إذا كان مجموع التكرارات فردى أو زوجي .

3 - تحديد الفترة ألتي يقع فيها الوسيط : وهى الفترة التي يقابلها تكرار متجمع صاعد يساوى رتبة الوسيط أو أكبر منها .

4 - تحديد طول الفترة الوسيطة وكذلك تكرارها .

5 - نستخدم القانون الآتي :

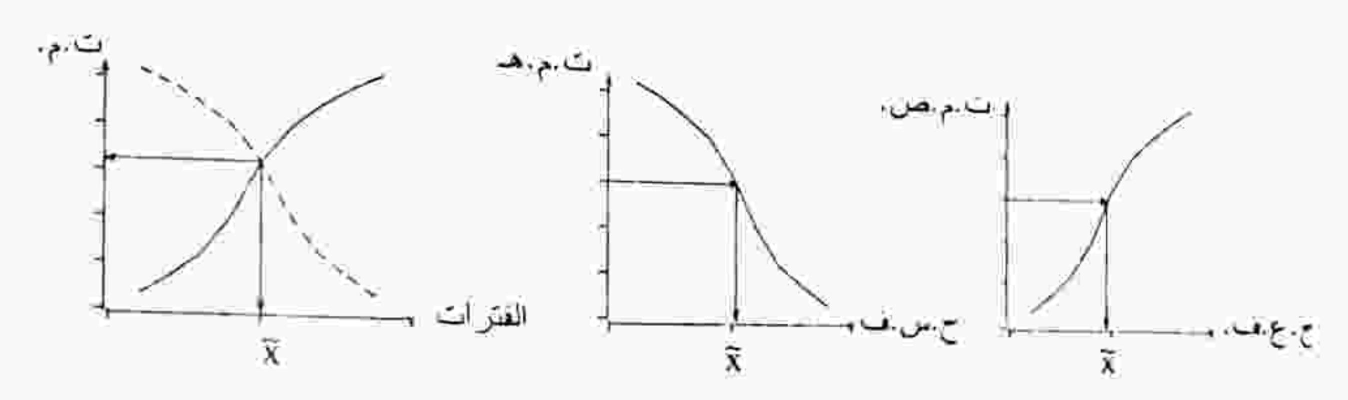
$$\tilde{x} = L + \left(\frac{\frac{n}{2} - F}{f}\right) \times c \quad (14)$$

حيث L = الحد الأدنى للفترة الوسيطة و f = تكرار الفترة الوسيطة و c = طول الفترة الوسيطة و F = التكرار المتجمع الصاعد للفترة التي تسبق الفترة الوسيطة . و $n = \sum_{i=1}^{k} f_i$

ألحظ أنه يجب أن يكون جدول التوزيع التكراري ذي فترات متصلة وإن لم يكن كذلك فيجب حساب الحدود الحقيقية للفترات .

(2) الطريقة البيانية

يمكن إيجاد الوسيط بيانياً باستخدام منحنى التكرار المتجمع الصاعد أو الهابط أو كلاهما معاً وذلك على النحو التالي :



شكل (10) :حساب الوسيط بيانياً

مثال (31) : أوجد الوسيط حسابياً وبيانياً مستخدماً بيانات المثال رقم (7)

الحـــل :

بإتباع الخطوات المشار إليها أعلاه نجد أن

الفترات	التكرار (f_i)	التكرار المتجمع الصاعد
56 - 50	1	1
62 - 56	5	6
68 - 62	12	18
74 - 68	15	33
80 - 74	22	55
86 - 80	11	66
92 - 86	6	72
98 - 92	3	75
المجموع	$\sum_{i=1}^{8} f_i = 75$	

رتبة الوسيط = $\frac{75}{2} = 37.5$ ،   الفترة الوسيطة : 74 - 80 .

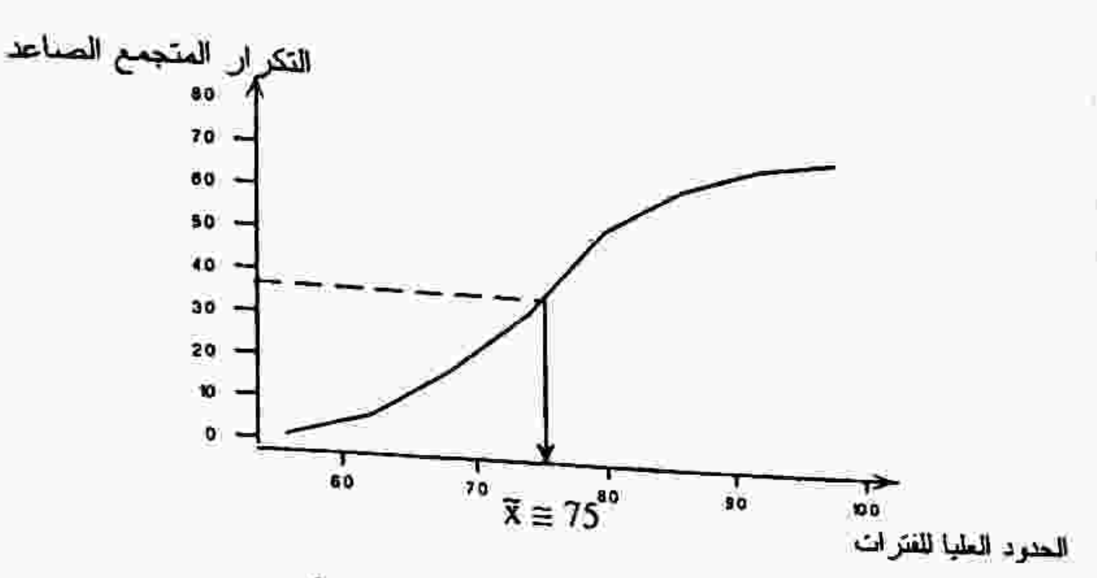
تكرارها (f) = 22 ،   طولها (c) = 80 - 74 = 6 .

التكرار المتجمع الصاعد للفترة التي تسبق فترة الوسيط (F) = 33 .

وعليه فإن

$$\tilde{x} = 74 + (\frac{37.5-33}{22}) \times 6$$
$$= 74 + (0.2045) \times 6 = 75.227$$

ويمكن حساب الوسيط بيانياً وذلك كما يلي :



شكل (11) : حساب الوسيط بيانياً

بعض خواص الوسيط :

1 - لا يتأثر بالقيم المتطرفة .

2 - يمكن حسابه بيانياً .

3 - يمكن حسابه من الجداول التكرارية المفتوحة .

4 - يمكن حسابه من البيانات الوصفية القابلة للترتيب وعددها فردى .

5 - لا يدخل في حسابه جميع القيم .

6 - لا نحتاج لتعديل التكرار إذا كانت أطوال الفترات غير متساوية .

7 - غير قابل للعمليات الجبرية .

8 - يتحدد بعدد البيانات وليس بقيمتها .

1 - 4 - 5 المنوال Mode

يستخدم هذا المقياس إذا كان الهدف من الدراسة الحصول على مقياس سريع لظاهرة النزعة المركزية بصرف النظر عن الدقة في القياس ، أو عندما يكون هناك اتجاه واضح نحو تمركز البيانات في جدول التوزيع التكراري ، ويستخدم المنوال كمقياس لوصف البيانات النوعية إلا أنه بالإمكان استخدامه أيضاً لوصف البيانات الكمية .

أ - طريقة حساب المنوال من البيانات الأولية

يتم حساب هذا المقياس في هذه الحالة وفقاً للتعريف الآتي :-

تعريف : المنوال هو القيمة أو الصفة الأكثر تكراراً مقارنة ببقية القيم أو الصفات .

إذن على ضوء التعريف المنوال لمجموعة من البيانات قد لا يوجد وإن وجد فقد لا يكون وحيداً وسوف نرمز للمنوال بالرمز m .

مثال (32) : أوجد المنوال لبيانات المثال رقم (17) .

الحـــل :

حيث أن المفردة 28 متكررة مرتين وهى الوحيدة المتكررة وعليه فإن المنوال $m = 28$.

ب – طريقة حساب المنوال من الجداول التكرارية

(1) الطريقة الحسابية :

لحساب المنوال من الجداول التكرارية نتبع الآتي :

1 – تحديد الفترة التي يقابلها أكبر تكرار أو أكبر تكرار معدل إذا كانت أطوال الفترات غير متساوي .

2 – حساب الفرق ما بين تكرار الفترة المنوالية وتكرار الفترة التي تسبقها ونرمز لذلك بالرمز Δ_1 ثم الفرق ما بين تكرار الفترة المنوالية وتكرار الفترة التي بعدها ويرمز لذلك بالرمز Δ_2 .

3 – تحديد الحد الأدنى للفترة المنوالية (L) وكذلك طولها (c) .

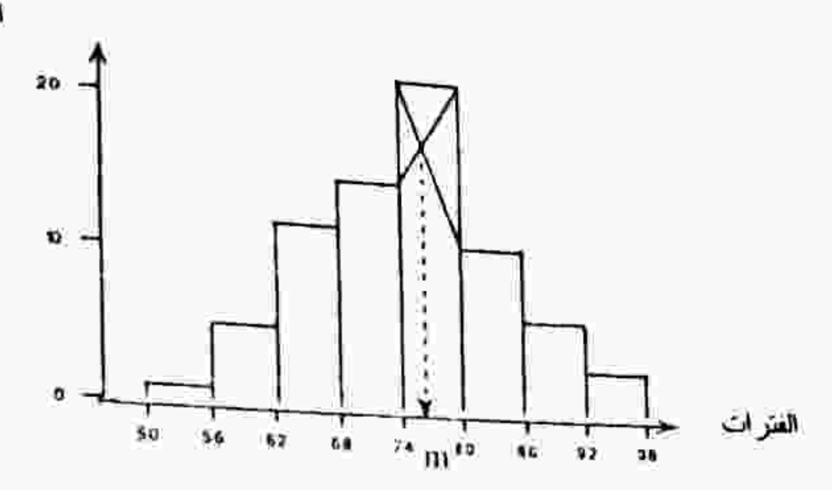
4 – نستخدم القانون الآتي لحساب المنوال :

$$m = L + \left(\frac{\Delta_1}{\Delta_1 + \Delta_2}\right) \times c \qquad (15)$$

ألحظ أنه يجب أن يكون جدول التوزيع التكراري ذي فترات متصلة وإن لم يكن كذلك فيجب استخدام الحدود الحقيقية للفترات .

(2) الطريقة البيانيـــة :

يمكن إيجاد المنوال بيانياً باستخدام المدرج التكراري حيث يتم توصيل نهاية المستطيل الذي يمثل الفترة المنوالية من الناحية اليسرى بنهاية المستطيل للفترة اللاحقة لها من الناحية اليسرى ، وكذلك نهاية المستطيل للفترة المنوالية من الناحية اليمنى بنهاية المستطيل للفترة السابقة لها من الناحية اليمنى ومن نقطة التقاطع نسقط عمود على المحور الأفقي ونقطة التقائه مع هذا المحور تعطى قيمة المنوال .



شكل (12) : حساب المنوال بيانياً

مثال (33) : أوجد المنوال حسابياً وبيانياً لبيانات المثال رقم (7) .

الحـــل :

الفترات	50-	56-	62-	68-	74-	80-	86-	98-92
التكرار (f)	1	5	12	15	22	11	6	3

حيث أن الفترة المنوالية هي الفترة التي يقابلها أكبر تكرار وعليه فإن

الفترة المنوالية : 74 - 80 ⇐ تكرارها = 22 وطولها (c) = 90 - 74 = 6 .

Δ_1 = تكرار الفترة المنوالية - تكرار الفترة التي تسبقها

= 22 - 15 = 7 .

Δ_2 = تكرار الفترة المنوالية - تكرار الفترة التي تليها

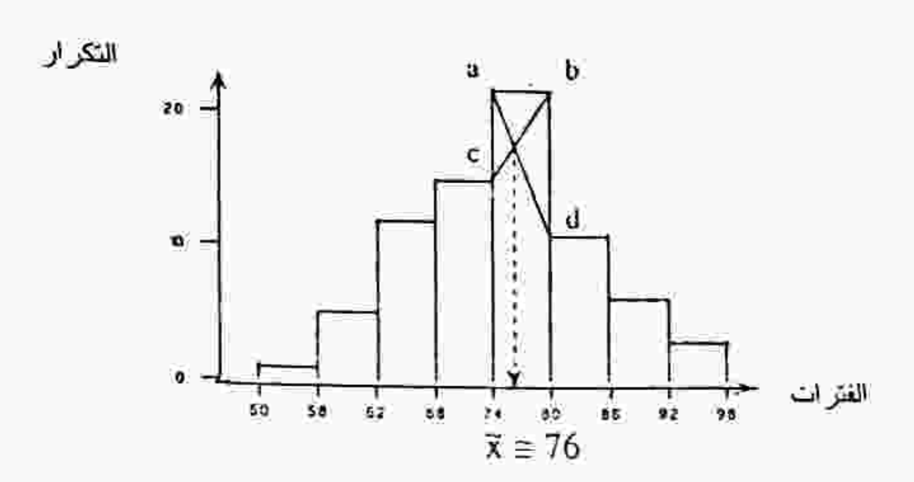
= 22 - 11 = 11 .

وعليه فإن المنوال سيكون كالآتي :

$$m = 74 + \left(\frac{7}{7+11}\right) \times 6$$

$$= 74 + \frac{42}{18} = 74 + 2.333 = 76.333$$

كما يمكن الحصول على هذه النتيجة بيانياً كالأتي :



شكل (13) : حساب المنوال بيانياً

إذا أنزلنا عمود على المحور الأفقي من نقطة تقاطع ad و bc فإن القيمة عند قاعدة هذا العمود تساوى المنوال (m) ، وبالتالي يكون المنوال مساوياً 76 تقريباً .

ملحوظة :--

يمكن حساب المنوال أيضاً بأحد الطرائق الأتية :

1 - تقريبياً : وذلك على اعتبار أنه يساوى مركز الفترة التي يقابلها أكبر تكرار فمن المثال السابق نجد أن

$$\text{المنوال} = \frac{80+74}{2} = 77$$

2 - طريقة الرافعة :- هذه الطريقة تنظر للفترة المنوالية على أنها ذراع أفقي تتجاذبه قوتين أحدهما تكرار القوة التي تسبق الفترة المنوالية والأخرى تكرار الفترة ما بعد الفترة المنوالية وبالتالي فإن قيمة المنوال تستقر في نقطة الفترة المنوالية ولكن سوف لن نتعرض لهذه الطريقة .
الحظ أن قيمة المنوال سوف تختلف قليلاً من طريقة إلى أخرى وذلك لأن كل طريقة مختلفة عن الأخرى . وبالتالي نفضل استخدام الطريقة الحسابية .

بعض خواص المنوال :

1 - أسهل مقاييس النزعة المركزية .

2 - يمكن حسابه بيانياً كما يمكن حسابه من الجداول التكرارية المفتوحة .

قليلة .

3 - ليس له معنى إذا كانت التكرارات قليلة .

4 - أفضل مقاييس النزعة المركزية لوصف الظواهر النوعية .

5- لا يتأثر بالقيم الشاذة .

6- يجب استخدام التكرار المعدل إذا كانت أطوال الفترات غير متساوية .

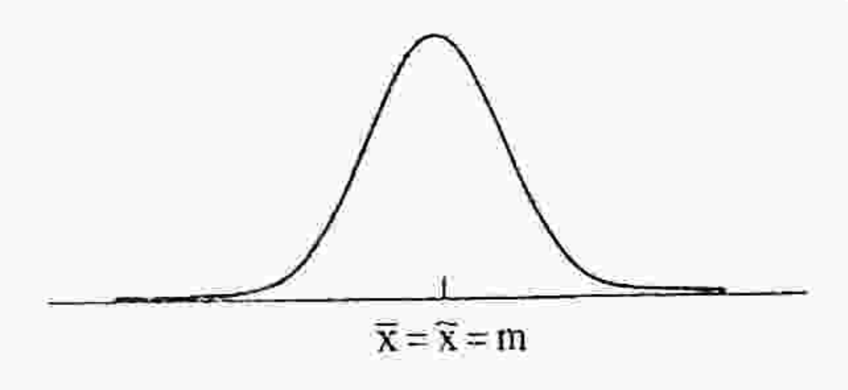
7 - لا يدخل في حسابه جميع البيانات المعطاة .

8 - غير قابل للعمليات الجبرية .

1 - 4 - 6 العلاقة بين المتوسط الحسابي والوسيط والمنوال

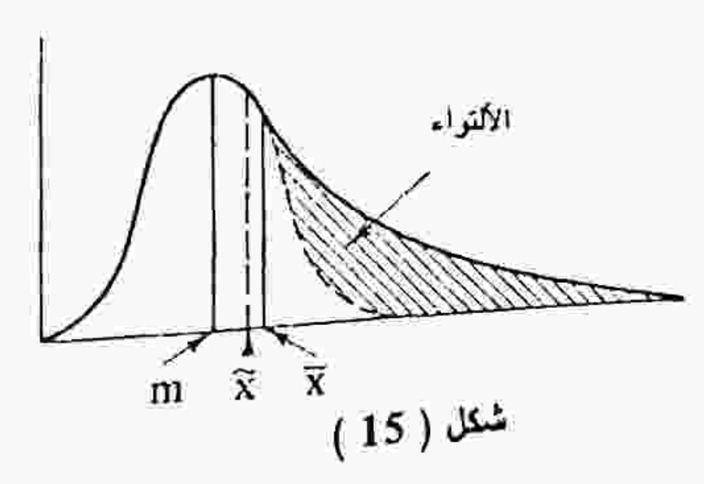
إذا كان لمجموعة البيانات منوال واحد فإن المتوسط والوسيط والمنوال تربطها إحدى العلاقات التالية :

أ - إذا كان التوزيع التكراري متماثل فإن المنحنى سيكون متماثل وله قمة واحدة وشكله يشبه شكل الجرس وفي هذه الحالة يكون المتوسط الحسابي مساوياً للوسيط ويساوى المنوال ويساوى مركز الفترة التي يقابلها أكبر تكرار . كما يتضح في الشكل التالي :



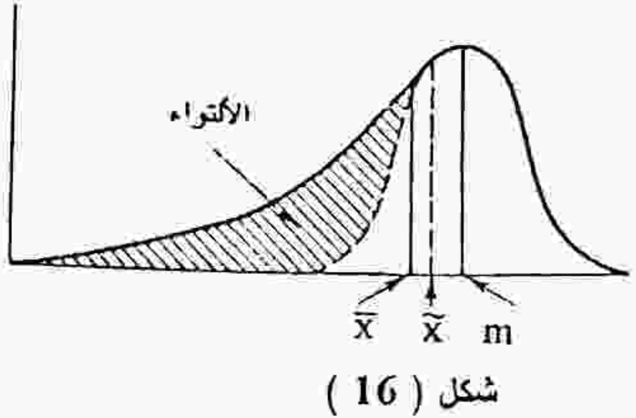
شكل (14)

ب - إذا كان المتوسط الحسابي أكبر من الوسيط والمنوال فإن المنحنى ملتويا ناحية اليمين . كما يتضح في الشكل التالي :



شكل (15)

جـ - إذا كان المتوسط الحسابي أقل من الوسيط والمنوال فإن المنحنى ملتويا ناحية اليسار . كما يتضح في الشكل التالي :



شكل (16)

وعليه إذا كان التوزيع التكراري وبالتالي المنحنى الذي يمثله ملتويا التواء قليلاً فإن العلاقة الآتية صحيحة

$$\bar{x} - m = 3(\bar{x} - \tilde{x}) \qquad (16)$$

وتفيد هذه العلاقة في إيجاد المتوسط الحسابي إذا كان جدول التوزيع التكراري مفتوح من أحد طرفيه .

1 - 4 - 7 الربيعيات والعشريات والمئويات

لحساب أي مقياس من هذه المقاييس نتبع نفس الخطوات التي أتبعناها عند حساب الوسيط حسابياً وبيانياً مع مراعاة الفرق في الرتبة فقط وإن لها نفس خواص الوسيط .

أ - الربيعيات Quartiles

1) - الربيع الأدنى Lower quartile (Q_1)

يعرف هذا المقياس على أنه القيمة أو المفردة التي يسبقها 25٪ من البيانات المرتبة ترتيباً تصاعدياً ويتم حسابه من الجداول التكرارية كما يلي :

$$Q_1 = L + \left(\frac{\frac{n}{4} - F}{f}\right) \times c \qquad (17)$$

حيث L = الحد الأدنى لفترة الربيع الأدنى و f = تكرار فترة الربيع الأدنى و c = طول فترة الربيع الأدنى و F = التكرار المتجمع الصاعد للفترة التي تسبق فترة الربيع الأدنى . و $n=\sum_{i=1}^{k} f_i$.

2) - الربيع الأعلى upper quartile (Q_3)

يعرف هذا المقياس على أنه القيمة أو المفردة التي يسبقها 75٪ من البيانات المرتبة ترتيباً تصاعدياً ويتم حسابه من الجداول التكرارية كما يلي :

$$Q_3 = L+(\frac{\frac{3n}{4} - F}{f})\times c \qquad (18)$$

حيث L = الحد الأدنى لفترة الربيع الأعلى و f = تكرار فترة الربيع الأعلى و c = طول فترة الربيع الأعلى و

F = التكرار المتجمع الصاعد للفترة التي تسبق فترة الربيع الأعلى . و $n=\sum_{i=1}^{k} f_i$

ب - العشريات Deciles (D)

يعرف العشير i حيث i = 1، 2 ، ....،9 على أنه القيمة أو المفردة التى يسبقها " 10 i ٪ " من البيانات المرتبة ترتيباً تصاعدياً ويتم حسابه من الجداول التكرارية كما يلي :

$$D_i = L+ (\frac{\frac{i\times n}{10} - F}{f})\times c \ , i=1,2,\cdots,9 \qquad (19)$$

حيث L = الحد الأدنى لفترة العشير i و f = تكرار فترة العشير i و c = طول فترة العشير i و

F = التكرار المتجمع الصاعد للفترة التى تسبق فترة العشير i . و $n=\sum_{i=1}^{k} f_i$

جـ - المئويات Percentiles (P)

يعرف المئوى i حيث i = 1، 2 ، ....،99 على أنه القيمة أو المفردة التى يسبقها " i ٪ " من البيانات المرتبة ترتيباً تصاعدياً ويتم حسابه من الجداول التكرارية كما يلي :

$$P_i = L + \left(\frac{\frac{i \times n}{100} - F}{f}\right) \times c \qquad , i = 1, 2, \cdots, 99 \qquad (20)$$

حيث L = الحد الأدنى لفترة المئوي i و f = تكرار فترة المئوي i و c = طول فترة المئوي i و

F = التكرار المتجمع الصاعد للفترة التي تسبق فترة المئوي i . و $n = \sum_{i=1}^{k} f_i$

ألحظ أنه يجب أن يكون جدول التوزيع التكراري ذي فترات متصلة وإن لم يكن كذلك فيجب حساب الحدود الحقيقية للفترات عند حساب أي مقياس من المقاييس أعلاه .وإنه إذا تساوت رتبة مقياسين أو أكثر فإن قيمها متساوية وعليه فإن :

(1) - $Q_1 = P_{25}$. (2) - $Q_3 = P_{75}$.

(3) - $D_5 = P_{50} = \tilde{x}$. (4) - $D_i = P_{10i}$.

مثال (34) : أوجد كلا من Q_1 و Q_3 و P_{10} و P_{90} .من البيانات التالية :

الفترات	10 -	20 -	30 -	40 -	50 -	60 -	70 -	80- 90
التكرار f_i	22	18	15	12	11	10	16	8

الحـل :

لحساب أي مقياس من هذه المقاييس وكما أشرنا سلفاً أننا نتبع نفس الخطوات التي أتبعناه عند حساب الوسيط ، وعليه سوف نوجد التكرار المتجمع الصاعد أولاً .

الفترات	التكرار f_i	التكرار المتجمع الصاعد
10 -	22	22
20 -	18	40
30 -	15	55
40 -	12	67
50 -	11	78
60 -	10	88
70 -	16	104
80 - 90	8	112
المجموع	112	

ألحظ أن رتبة الربيع i تساوى $\frac{i \times n}{4}$ حيث i = 1 ، 2 ، 3 .

رتبة العشير i تساوى $\frac{i \times n}{10}$ حيث i = 1 ، 2 ، ... ، 9 .

رتبة المئوى i تساوى $\frac{i \times n}{100}$ حيث i = 1 ، 2 ، ... ، 99 . وعليه فإن

1 - الربيع الأدنى :

رتبة الربيع الأدنى $= \frac{\sum_{i=1}^{8} f_i}{4} = \frac{112}{4} = 28$ وعليه فإن فترة الربيع الأدنى هي 20 - 30

طولها= 30 - 20 =10 ، تكرارها= 18 ، التكرار المتجمع الذي يسبق فترة الربيع = 22 ، وبالتالي فإن

$$Q_1 = L + (\frac{\frac{n}{4} - F}{f}) \times c = 20 + (\frac{28-22}{18}) \times (10) = 23.3333$$

2- الربيع الأعلى :

رتبة الربيع الأعلى $= \frac{3 \times \sum_{i=1}^{8} f_i}{4} = \frac{3 \times 112}{4} = 84$ ، وعليه فإن فترة الربيع الأعلى هي 60 - 70 و طولها = 70 - 60 =10 ، تكرارها= 10 ، التكرار المتجمع الذي يسبق فترة الربيع = 78 ، وبالتالي فإن

$$Q_3 = L + (\frac{\frac{3n}{4} - F}{f}) \times c = 60 + (\frac{84-78}{10}) \times (10) = 66$$

3- المئوي العاشر :

رتبة المئوي العاشر $= \frac{10 \times \sum_{i=1}^{8} f_i}{100} = \frac{10 \times 112}{100} = 11.2$ وعليه فإن فترة المئوي العاشر هي 10 - 20

طولها= 20 - 10 =10 ، تكرارها= 22 ، التكرار المتجمع الذي يسبق فترة المئوي = 0 ، وبالتالي فإن

$$P_{10} = L + (\frac{\frac{i \times n}{100} - F}{f}) \times c = 10 + (\frac{11.2-0}{22}) \times (10) = 15.0909$$

4 - المئوي تسعون :

رتبة المئوي تسعون = $\frac{90 \times \sum_{i=1}^{8} f_i}{100} = \frac{90 \times 112}{100} = 100.8$ وعليه فإن فترة المئوي تسعون هي 70 - 80 ، طولها= 80 - 70 =10 ، تكرارها= 16 ، التكرار المتجمع الذي يسبق فترة المئوي =88 وبالتالي فإن :

$$P_{90} = L + \left(\frac{\frac{i \times n}{100} - F}{f}\right) \times c = 70 + \left(\frac{100.8 - 88}{16}\right) \times (10) = 78$$

ألحظ أنه يمكن حساب هذه المقاييس بيانياً كما في حالة الوسيط .

1 - 5 مقاييس التشتت Measures of Dispersion (Variation)

لقد تعرضنا في البنود السابقة إلى كيفية تلخيص عدد كبير من البيانات الإحصائية بطرق تمكننا من فهمها وتحليلها من الناحية الإحصائية ، ثم كيفية البحث عن قيمة واحدة يمكن أن توصف بها هذه البيانات نعتقد بأن قيم الظاهرة أو "المتغير موضوع الدراسة تميل للتمركز حولها، ولكن أي متوسط لوحده لا يكفي لقياس هذا الاتجاه نحو التمركز ، وبالتالي يفضل وجود مقاييس أخرى نستطيع من خلالها وصف التوزيع التكراري وصفاً كاملاً ومقارنته بأي توزيع تكراري آخر . هذه المقاييس يمكن تقسيمها إلى ثلاث وهي :-

أ - مقياس يقيس مدى تباعد أو تمركز القيم حول القيمة المتوسطة أي مدى اختلافها أو تشتتها .

ب - مقياس يقيس انحراف القيم عن التماثل حول القيمة المتوسطة (الالتواء) .

جـ - مقياس يقيس درجة تجمع القيم عند القيمة المتوالية (التفرطح) .

تكمن أهمية مقياس التشتت في كون أنه لا يمكن أن نتصور مثلاً تساوي الإنتاج في جميع المؤسسات الصناعية أو تساوي مستوى الخدمات في جميع المواقع الخدمية أو تساوي أطوال جميع الطلاب ... الخ ، وبالتالي استخدام قيمة واحدة لوصف التوزيع التكراري قد تكون مضللة أحياناً ، فمثلاً إذا كان لدينا المجموعتان التاليتان من القيم :-

(1) - 7 ، 0 ، 14 .   (2) - 7 ، 6 ، 8 .

فإن المتوسط الحسابي لكل منهما يساوي 7 وإذا اكتفينا بهذا المقياس فإننا نقرر أن المجموعتين متشابهتان ، ولكن في الحقيقة أن قيم المجموعة الأولى أكثر تباعداً من قيم المجموعة الثانية ، وهنا يأتي دور مقاييس التشتت أو الاختلاف لوصف هذه الناحية في البيانات الإحصائية .

ويمكن تقسيم مقاييس التشتت إلى مجموعتين هما : مقاييس تقيس تقارب أو تباعد القيم عن بعضها البعض وهى : المدى والانحراف الربيعي (أو نصف المدى الربيعي) ، والأخرى مقاييس تقيس قرب أو بعد القيم من قيمة معينة كالمتوسط الحسابي مثلاً وهى : الانحراف المتوسط المطلق والانحراف المعياري .

1 - 5 - 1 المدى Range (R)

يستخدم هذا المقياس عندما يكون الهدف هو الحصول على مقياس سريع لمدى تشتت المفردات دون الاهتمام الكبير بالدقة في القياس أو حين ما يكون للمفردات المتطرفة أهمية خاصة .

أ - طريقة حساب المدى من البيانات الأولية

يتم حساب المدى في هذه الحالة كما يلي :

$$R = x_{(n)} - x_{(1)} \qquad (21)$$

حيث $x_{(1)}$ = أصغر مفردة و $x_{(n)}$ = أكبر مفردة .

مثال (35): من بيانات المثال رقم (17) أوجد المدى .

الحـــل :

حيث أن أكبر قيمة في البيانات تساوى 30 وأصغر قيمة هي 20 وعليه فإن المدى :

$$R = x_{(n)} - x_{(1)} = 30 - 20 = 10$$

ب - طريقة حساب المدى من الجداول التكرارية

يتم حساب المدى في هذه الحالة بإحدى الطريقتين الآتيتين :-

1 - $R = U_k - L_1$

حيث U_k = نهاية الفترة الأخيرة ، L_1 = بداية الفترة الأولى .

2 - $R = M_k - M_1$

حيث M_k = مركز الفترة الأخيرة ، M_1 = مركز الفترة الأولى .

إن الطريقة (2) تبدو أنها تزيل أثر القيم المتطرفة .

مثال (36) : أوجد المدى مستخدماً بيانات المثال رقم (7) .

الحــل :

مــن الصيغــة (1) نجــد أن $R=98-50=48$ ومــن الصيغــة (2) نجــد أن $R=95-53=42$.

ألحظ أن الناتج يختلف وذلك لأن كل طريقة قائمة على أساس مختلف .

بعض خواص المدى :

1 - إذا كانت جميع المفردات متساوية فإن المدى يساوى صفر آي أنه لا يوجد تشتت .

2 - مقياس مضلل في حالة وجود قيم شادة .

3 - لا يمكن حسابه من الجداول التكرارية المفتوحة .

4 - يستخدم في رسم الخرائط الإحصائية لمراقبة مطابقة الإنتاج للمواصفات المطلوبة .

5 - يمكن استخدامه للمقارنة بين مجموعتين مختلفتين من البيانات إذا اختلفت في المتوسطات الحسابية وتساوت في المدى وهو يسمى بالمدى النسبي، في هذه الحالة ، حيث المدى النسبي = $\frac{R}{\bar{x}} \times 100$.

6 - بسيط الحساب وسهل المفهوم ولذلك فهو كثير الاستخدام في الأوساط العامة .

7 - لا يعتمد في حسابه على كل البيانات .

8 - هو عبارة عن فترة تحتوى على كل البيانات .

مثال (37): إذا كان متوسط الدخل اليومي لمصحة سكره خلال فترة زمنية معينة يساوى 250 دينار وبمدى 10 بينما متوسط ومدى الدخل اليومي لمصحة العافية خلال نفس الفترة يساوى 300 و 10 على التوالي فأوجد المدى النسبي لدخل المصحتين .

الحــل :

المدى النسبي لدخل لمصحة سكره = $\frac{10}{250} \times 100 = 4\%$.

المدى النسبي لدخل لمصحة العافية = $\frac{10}{300} \times 100 = 3.33\%$.

حيث أن المدى النسبي لدخل مصحة العافية أقل من المدى النسبي لدخل مصحة سكره اليومي، وعليه فإن الدخل اليومي لمصحة العافية أكثر انتظاماً آي أقل تشتتاً (أكثر تجانساً) من مصحة سكره .

1 - 5 - 2 الانحراف الربيعي (Quartile Devaiation (Q. D.

للتخلص من بعض عيوب المدى والتي من أهمها تأثره بالقيم الشاذة وعدم إمكانية حسابه في حالة الجداول التكرارية المفتوحة وجد مقياس آخر لظاهرة التشتت وهو ما يسمى بالانحراف الربيعي والفكرة الأساسية في هذا المقياس هي إهمال الربع الأول والأخير من البيانات المرتبة ترتيباً تصاعديا ، وفي مثل هذه الحالة تكون أكبر قيمة في البيانات هي الربيع الأعلى وأصغر قيمة الربيع الأدنى والفرق بينهما يعطى ما يسمى بالمدى الربيعي وهو ما يطلق عليه تسمية الانحراف الربيعي ، وبقسمة المدى الربيعي على 2 نتحصل على ما يسمى بنصف المدى الربيعي. ويستخدم هذا المقياس إذا كان الوسيط هو المقياس المناسب للنزعة المركزية أو عندما يكون جدول التوزيع التكراري مفتوحاً أو شديد الالتواء أو عندما يكون هناك قيم متطرفة .

ويتم حساب هذا المقياس سواء من البيانات الأولية أو الجداول التكرارية كما يلي :

$$Q.D = Q_3 - Q_1 \qquad (22)$$

مثال (38) : من بيانات مثال (34) أوجد الانحراف الربيعي .

الحـــل :

حيث أن $Q_1 = 23.3333$ و $Q_2 = 66$ ، وعليه فإن :

$$Q.D. = Q_3 - Q_1 = 66 - 23.3333 = 42.6667$$

بعض خواص الانحراف الربيعي :

1 - لا يتأثر بالقيم الشاذة (المتطرفة) .

2 - يمكن الاستعانة بالطريقة البيانية في حسابه .

3 - يستخدم كمقياس للتشتت في التوزيعات التكرارية شديدة الالتواء .

4 - يمكن حسابه في حالة الجداول التكرارية المفتوحة .

5 - يتحدد بعدد البيانات وليس بقيمتها .

6 - هو عبارة عن فترة تحتوى على 50 ٪ من البيانات .

1 - 5 - 3 الانحراف المتوسط Mean Deviation (M. D.)

يعتمد كل من المدى والانحراف الربيعي على قيمتين فقط فالأول يعتمد على أكبر وأصغر قيمة في البيانات بينما يعتمد الثاني على قيمتي الربيع الأدنى والربيع الأعلى ، لذلك وجد مقياس أخر يعتمد في حسابه على كل القيم . والفكرة الأساسية في هذا المقياس هي قياس مدى تباعد (انحراف) القيم عن متوسطها الحسابي بغض النظر فيما إذا كان ذلك الانحراف سلبياً أو إيجابياً. فكلما كانت تلك القيم قريبة من متوسطها دل ذلك على تجانسها والعكس صحيح .

أ - حساب الانحراف المتوسط في حالة البيانات الأولية

إذا كانت البيانات $x_1, x_2, x_3, \ldots, x_n$ تمثل قيم لظاهرة ما فإن الانحراف المتوسط لهذه البيانات سنرمز له بالرمز .M.D ويتم حسابه كما يلي :

$$M.D. = \frac{\sum_{i=1}^{n} |x_i - \bar{x}|}{n} \qquad (23)$$

مثال (39) : أوجد الانحراف المتوسط للبيانات التالية :

10 ، 7 ، 5 ، 3 ، 2

الحـل :

يمكن توضيح خطوات الحل في الجدول التالي :

x_i	$x_i - \bar{x}$	$\lvert x_i - \bar{x} \rvert$
2	− 3.4	3.4
3	− 2.4	2.4
5	− 0.4	0.4
7	1.6	1.6
10	4.6	4.6
27	0	12.4

حيث $\bar{x}=\frac{\sum_{i=1}^{n} x_i}{n}=\frac{27}{5}=5.4$ وعليه فإن

$$M.D.=\frac{\sum |x_i-\bar{x}|}{n}=\frac{12.4}{5}=2.48$$

ب- طريقة حساب الانحراف المتوسط من الجداول التكرارية

إذا كانت $x_1, x_2, x_3, \dots, x_k$ تمثل المفردات المختلفة للظاهرة x (أو مراكز الفترات) و $f_1, f_2, f_3, \dots, f_k$ تمثل التكرارات المناظرة لها فإن الانحراف المتوسط لهذه البيانات يُعرف كما يلي :

$$M.D.=\frac{\sum_{i=1}^{k} f_i |x_i-\bar{x}|}{n} \qquad (24)$$

حيث $\bar{x}=\frac{\sum_{i=1}^{k} f_i x_i}{n}$.

مثال (40) : أوجد الانحراف المتوسط للبيانات التالية :

الفترات	0 -	2 -	4 -	6 -	8 - 10	المجموع
التكرار	2	5	4	8	1	20

الحـــل :

الفترات	التكرار f_i	مراكز الفترات x_i	$f_i x_i$	$\lvert x_i-\bar{x}\rvert$	$f_i \lvert x_i-\bar{x}\rvert$
0 -	2	1	2	4.1	8.2
2 -	5	3	15	2.1	10.5
4 -	4	5	20	0.1	0.4
6 -	8	7	56	1.9	15.2
8 - 10	1	9	9	3.9	3.9
المجموع	20		102		38.2

حيث $\bar{x} = \frac{\sum_{i=1}^{n} f_i x_i}{n} = \frac{102}{20} = 5.1$ وعليه فإن

$$M.D. = \frac{\sum_{i=1}^{n} f_i |x_i - \bar{x}|}{n} = \frac{38.2}{20} = 1.91$$

بعض خواص الانحراف المتوسط :

1 - يتأثر بالقيم الشاذة .

2 - لا يمكن حسابه في حالة الجداول التكرارية المفتوحة .

3 - يعتمد في حسابه على كل القيم .

4 - قليل الاستخدام في الإحصاء الاستنتاجي لعدم قابليته للعمليات الجبرية .

5 - يمكن حسابه عن طريق الانحرافات عن الوسيط مع الملاحظة أن الانحراف المتوسط عن الوسيط أقل من الانحراف المتوسط عن المتوسط الحسابي .

1 - 5 - 4 الانحراف المعياري Standard Deviation (S.D.)

مما سبق يتضح أن الانحراف المتوسط أفضل من كل من المدى والانحراف الربيعي وذلك لأنه يعتمد في حسابه على كل البيانات، ولكن يؤخذ على هذا المقياس عدم قابليته للعمليات الجبرية، لذلك وجد مقياس آخر لظاهرة التشتت وهو ما يسمى بالانحراف المعياري ويعرف على أنه الجذر التربيعي للتباين الذي يرمز له بالرمز S^2 .

أ - حساب الانحراف المعياري في حالة البيانات الأولية

إذا كانت البيانات $x_1 , x_2 , x_3 , \ldots , x_n$ تمثل قيم لظاهرة ما فإن الانحراف المعياري لهذه البيانات سيرمز له بالرمز S.D ويتم حسابه كما يلي :

$$S.D = \sqrt{S^2} \qquad (25)$$

حيث التباين معرف كالتالي :

$$S^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1} \qquad (26)$$

ويمكن إعادة كتابة ذلك بطريقة أخرى (الصيغة العملية) على النحو التالي :

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left[x_i^2 - 2\bar{x}x_i + \bar{x}^2\right]$$

$$= \frac{1}{n-1}\left[\sum_{i=1}^{n}x_i^2 - 2\bar{x}\sum_{i=1}^{n}x_i + \sum_{i=1}^{n}\bar{x}^2\right]$$

$$= \frac{1}{n-1}\left[\sum_{i=1}^{n}x_i^2 - 2n\bar{x}^2 + n\bar{x}^2\right] = \frac{1}{n-1}\left[\sum_{i=1}^{n}x_i^2 - n\bar{x}^2\right] \qquad (27)$$

مثال (41): أوجد الانحراف المعياري مستخدماً بيانات المثال رقم(17) .

الحــل :

حيث أن

$$\bar{x} = \frac{\sum_{i=1}^{6}x_i}{n} = \frac{30+25+20+26+28+24+28+27}{8} = \frac{208}{8} = 26$$

وإن

$$\sum_{i=1}^{6}x_i^2 = (30)^2+(25)^2+(20)^2+(26)^2+(28)^2+(24)^2+(28)^2+(27)^2$$

$$= 900+625+400+676+784+576+784+729 = 5474$$

وعليه فإن

$$s^2 = \frac{1}{n-1}\left[\sum_{i=1}^{n}x_i^2 - n\bar{x}^2\right] = \frac{1}{7}[5474-(8)(26)^2] = \frac{66}{7} = 9.429$$

$$S.D. = \sqrt{9.429} = 3.071$$

ب - طريقة حساب الانحراف المعياري من الجداول التكرارية

إذا كانت $x_1 , x_2 , x_3 , \ldots , x_k$ تمثل المفردات المختلفة للظاهرة x (أو مراكز الفترات) و $f_1 , f_2 , f_3 , \ldots , f_k$ تمثل التكرارات المناظرة لها فإن الانحراف المعياري لهذه البيانات يُعرف كما يلي :

$$S.D. = \sqrt{s^2}$$

حيث :

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{k} f_i (x_i - \overline{x})^2 = \frac{1}{n-1}\left[\sum_{i=1}^{k} f_i x_i^2 - n\,\overline{x}^2\right] \qquad (28)$$

و $\overline{x} = \dfrac{\sum_{i=1}^{k} f_i x_i}{n}$ و $n = \sum_{i=1}^{k} f_i$.

مثال (42) : أوجد التباين والانحراف المعياري مستخدماً بيانات المثال رقم (7) .

الحــل :

الفترات	التكرار f_i	مراكز الفترات x_i	$f_i x_i$	$f_i x_i^2$
56-50	1	53	53	2809
62-56	5	59	295	17405
68-62	12	65	780	50700
74-68	15	71	1065	75615
80-74	22	77	1694	130438
86-80	11	83	913	75779
92-86	6	89	534	47526
98-92	3	95	285	27075
المجموع	$\sum_{i=1}^{8} f_i = 75$		5619	$\sum_{i=1}^{8} f_i x_i^2 = 427347$

$$\bar{x} = \frac{\sum_{i=1}^{n} f_i\, x_i}{n} = \frac{5619}{75} = 74.92$$

$$s^2 = \frac{1}{n-1}\left[\sum_{i=1}^{k} f_i\, x_i^2 - n\,\bar{x}^2\right] = \frac{1}{74}[\,427347 - 75(74.92)^2\,]$$

$$= \frac{6371.52}{74} = 86.1016$$

وعليه فإن

$$S.D. = \sqrt{86.1016} = 9.2791$$

بعض خواص الانحراف المعياري :

1- لا يمكن حسابه من الجداول التكرارية المفتوحة .

2- يتأثر بالقيم المتطرفة وذلك لأنه يعتمد في حسابه على المتوسط الحسابي الذي بدوره يتأثر بها.

3- قابل للعمليات الجبرية ، ولذلك فهو كثير الاستخدام في القوانين والنظريات الإحصائية .

4- إذا كانت الظاهرة X تأخذ القيم $x_1, x_2, x_3, \cdots, x_m$ و $\bar{x}$ تمثل المتوسط الحسابي و s_1^2 تمثل التباين لها وكانت الظاهرة Y تأخذ القيم $y_1, y_2, \cdots, y_n$ و $\bar{y}$ تمثل المتوسط الحسابي و s_2^2 تمثل التباين لها فإن التباين المشترك للظاهرتين معاً نرمز له بالرمز s_p^2 (Pooled Variance) ومعرف كما يلي :

$$s_p^2 = \frac{1}{m+n}\{m s_1^2 + n s_2^2 + \tfrac{mn}{m+n}(\bar{x}-\bar{y})^2\} \tag{29}$$

وإذا كانت $\bar{x} = \bar{y}$ فإن

$$s_p^2 = \frac{1}{m+n}\{m s_1^2 + n s_2^2\} \tag{30}$$

حيث $s_1^2 = \dfrac{\sum_{i=1}^{m}(x_i-\bar{x})^2}{m}$ و $s_2^2 = \dfrac{\sum_{i=1}^{n}(y_i-\bar{y})^2}{n}$ وغالباً ما نستخدم الصيغة التالية في حساب التباين المشترك :

$$s_p^2 = \frac{(m-1)s_1^2 + (n-1)s_2^2}{m+n-2} \tag{31}$$

حيث المقام في s_1^2 مقسوماً على m-1 وفي حالة s_2^2 مقسوماً على n-1 بدلاً من m و n .
5 - يعد أهم مقاييس التشتت وأكثرها استخداماً .

1- 5 - 5 معامل الاختلاف Coefficient of variation (c. v.)

إن مقاييس التشتت السابقة تعتمد جميعها على الوحدات المستخدمة في القياس وبالتالي لا يمكن استخدامها في المقارنة بين مجموعتين أو أكثر مقاسه بوحدات قياس مختلفة مثل الأطوال والأوزان والدرجات مثلاً . ومن ناحية أخرى لو أردنا المقارنة من حيث التشتت حتماً سيتم استخدام أفضل تلك المقاييس وهو الانحراف المعياري ، وحيث أن الانحراف المعياري يقيس مقدار انحراف القيم عن متوسطها وعليه فإن عملية المقارنة تكون غير واقعية إذا اختلفت المتوسطات الحسابية للمجموعات المراد مقارنتها حتى ولو كانت مقاسه بنفس الوحدات . لذلك وجدت مقاييس أخرى لا تعتمد على الوحدات المستخدمة في القياس حيث تقيس الاختلاف النسبي دون وحدة تمييز أهمها وأكثرها استخداماً معامل الاختلاف وهو عبارة عن النسبة المئوية للانحراف المعياري منسوباً إلى المتوسط الحسابي وبالطبع كلما كانت قيمة هذا المعامل كبيرة دل ذلك على وجود تشتت كبير بين مفردات التوزيع والعكس صحيح . ويرمز لهذا المعامل بالرمز c . v ومعرف كما يلي :

$$c.v = \frac{s.d}{\bar{x}} \times 100 \qquad (32)$$

مثال (43) : إذا علمت أن متوسط درجة الحرارة في مدينة طرابلس في شهر مايو السنة الماضية كان 24 درجة مئوية وبانحراف معياري 3 درجات مئوية ، بينما متوسط درجة الحرارة في نفس الشهر في السنة الماضية في مدينة سبها كان 34 درجة مئوية وبانحراف معياري 4 درجات مئوية فأي المدينتين أقل تشتتاً (أكثر تجانساً) من حيث درجة الحرارة .

الحـــل :

معامل الاختلاف لدرجات الحرارة في مدينة طرابلس :

$$c.v = \frac{s.d}{\bar{x}} \times 100 = \frac{3}{24} \times 100 = 12.5\%$$

معامل الاختلاف لدرجات الحرارة في مدينة سبها :

$$c.v = \frac{s.d}{\bar{x}} \times 100 = \frac{4}{34} \times 100 = 11.77\%$$

وعليه فإن درجات الحرارة في مدينة سبها أكثر تجانساً (أقل اختلافاً) من درجات الحرارة في مدينة طرابلس .

بعض خواص معامل الاختلاف :

1- يقيس الاختلاف النسبي دون وحدة تمييز .

2- ليس له معنى إذا كانت المتوسطات الحسابية تساوى صفراً .

3- يستخدم في مقارنة مجموعتين أو أكثر من البيانات من حيث التشتت ، وخاصة إذا اختلفت المتوسطات الحسابية .

1 - 5 - 6 العزوم Moments

إذا كانت الظاهرة X تأخذ القيم $x_1 , x_2 , x_3 , \cdots , x_m$ و $\bar{x}$ تمثل المتوسط الحسابي لها فإن العزم الرائي حول المتوسط الحسابي في حالة البيانات الأولية معرف كما يلي :

$$m'_r = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^r}{n} \quad , \ r = 1, 2, 3, 4, \cdots \qquad (33)$$

وعليه إذا كانت

$$r = 1 \Rightarrow m'_1 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})}{n} = 0$$

$$r = 2 \Rightarrow m'_2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n}$$

$$\vdots$$

... وهكذا .

وإذا كانت $\bar{x} = 0$ فإن العزم الرائي حول نقطة الأصل (الصفر) يكون كالآتي :

$$m_r = \frac{\sum_{i=1}^{n} x_i^r}{n} \quad , \ r = 1, 2, 3, 4, \cdots \qquad (34)$$

ويمكن تعريف العزم حول أي قيمة أخرى خلاف المتوسط الحسابي وذلك من خلال استبدال المتوسط الحسابي بتلك القيمة في الصيغة أعلاه .أما في حالة الجداول التكرارية فإنه يتم حساب العزم الرائي حول المتوسط الحسابي كما يلي :

$$m'_r = \frac{\sum_{i=1}^{k} f_i (x_i - \bar{x})^r}{n} \quad , \; r=1,2,3,4,\cdots \qquad (35)$$

وإذا كانت $\bar{x}=0$ فإن :

$$m_r = \frac{\sum_{i=1}^{k} f_i \; x_i^r}{n} \quad , \; r=1,2,3,4,\cdots \qquad (36)$$

1 - 5 - 7 الالتواء والتفرطح Skewness and Kurtosis

الالتواء : يعبر الالتواء (Skewness) عن درجة توزيع البيانات حول نقطة التمركز فيها ، فوجود الالتواء دليل على انعدام الانتظام في التوزيع . ويمكن معرفة طبيعة ودرجة التواء أي توزيع بمجرد النظر إلى شكله البياني ، ولكن كثيراً ما نحتاج لتقدير درجة الالتواء بدقة وبالتالي يجب استخدام مقياس دقيق لهذه الظاهرة الهامة ، وسنقسم مقاييس الالتواء إلى نوعين أساسيين هما :-

أ- إذا كان المتوسط الحسابي يمثل نقطة التمركز في التوزيع :

إذا كان التوزيع التكراري متماثلاً فإن المتوسط الحسابي والوسيط والمنوال تتساوى جميعها عند النقطة المقابلة لقمة المنحنى وكلما بعد التوزيع عن التماثل كلما اختلفت هذه المتوسطات الثلاث عن بعضها ، وبذلك أقترح بيرسون مقاييس للالتواء عرفه كالآتي :-

$$\alpha_1 = \frac{\bar{x} - m}{s.d.} \qquad (37)$$

حيث $\bar{x}$: المتوسط الحسابي و m : المنوال و s.d. : الانحراف المعياري .

وعليه إذا كان معامل الالتواء يساوى صفراً فإن المنحنى متماثل ، وإذا كان معامل الالتواء أكبر من الصفر فإن الالتواء موجب ، وإذا كان معامل الالتواء أقل من الصفر فإن الالتواء سالب .

ولكن المقياس السابق يعتمد على المنوال وهو مقياس غير دقيق لا يجب الاعتماد عليه في قياس الالتواء ، وبالتالي يفضل استخدام العلاقة التالية في حالة التوزيعات القريبة من التماثل (المعتدلة) وهى كالآتي :

$$\alpha_2 = \frac{3(\bar{x} - \tilde{x})}{s.d.} \quad .$$

(38)

مثال (44): إذا علمت أن $\bar{x} = 74.92$ و $m = 76.33$ و $\tilde{x} = 75.227$ $s = 22.16996$

أوجد كلا من α_1 و α_2 .

الحـل :

من التعريف أعلاه نجد أن

$$\alpha_1 = \frac{\bar{x} - m}{s.d.} = \frac{74.92 - 76.333}{22.16996} = -0.0637$$

و

$$\alpha_2 = \frac{3(\bar{x} - \tilde{x})}{s.d.} = \frac{3(74.92 - 75.227)}{22.16996} = -0.0415$$

وحيث أن $\alpha_1 < 0$ و $\alpha_2 < 0$ وعليه فإن منحنى هذه البيانات ملتويا ناحية اليسار أي أن الالتواء سالب .

الحظ أن القيمة المطلقة للمقياسين مختلفة بالرغم من أن الخلاصة واحدة ، وذلك لأن كل منها قائم على أساس مختلف .

ب - إذا كان الوسيط يمثل نقطة التمركز في التوزيع :

إذا كان التوزيع مفتوحاً فإن الالتواء يقاس هنا بمقياس قائم على أساس العلاقة بين الربيع الأعلى والربيع الأدنى والوسيط ، وذلك لأنه في حالة التوزيع التكراري المتماثل تتساوى المسافة بين كل من الربيعين والوسيط ولا يتحقق ذلك إذا كان التوزيع ملتوياً ويعرف هذا المقياس كما يلي:

$$\alpha_3 = \frac{(Q_3 - \tilde{x}) - (\tilde{x} - Q_1)}{Q_3 - Q_1} \qquad (39)$$

ويطلق على هذا المقياس أحياناً تسمية معامل الالتواء الربيعي .

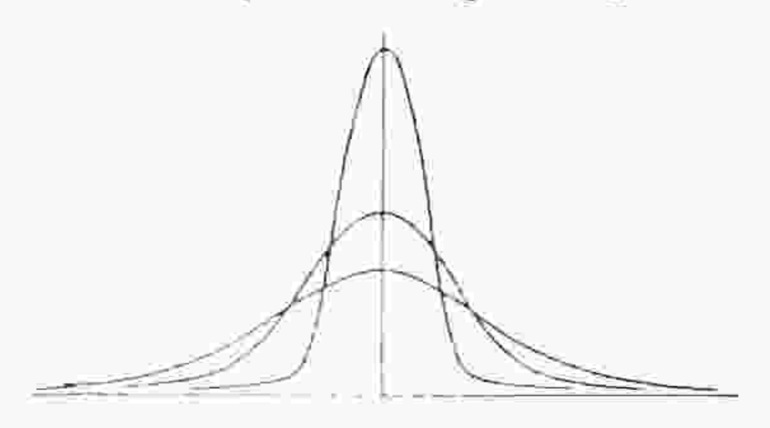
ويجب التذكير هنا أنه إذا أردنا المقارنة بين درجة التواء توزيعين تكراريين أو أكثر يجب استخدام نفس المقياس فلا يجوز استخدام مقاييس مختلفة وذلك لأن كل منها وكما أشرنا سلفاً قائم على أسس مختلفة .

التفرطح : يعبر التفرطح (Kurtosis) عن درجة تدبب المنحنى التكراري بالنسبة إلى التوزيع المعتدل (المتماثل) ، فإذا كانت البيانات أكثر تجمعاً حول المنوال أي أن قاعدة التوزيع ضيقة وطرفاه مرتفعين فإنه يقال بأن التوزيع مدبب ، أما إذا كانت القيم كثيرة على الجانبين أي أن قاعدة المنحنى واسعة وطرفاه منخفضين فإنه يقال بأن المنحنى متفرطح ويقاس التفرطح كما يلي :

$$\gamma = \frac{m'_4}{(s.d.)^4} \qquad (40)$$

حيث m'_4 تمثل العزم الرابع حول المتوسط الحسابي .

وعليه إذا كانت $\gamma = 3$ فإن المنحنى معتدل التفرطح ، وإذا كانت $\gamma > 3$ فإن المنحنى مدبب ، أما إذا كانت $\gamma < 3$ فإن المنحنى متفرطح كما يتضح من الشكل التالي :-



شكل (17) : منحنيات تكرارية متمركزة حول نفس النقطة ولكنها مختلفة في تدببها .

ويمكن قياس التفرطح بمقياس أخر يسمى بمعامل التفرطح المئوي (β) وهو معرف كما يلي :

$$\beta = \frac{Q_3 - Q_1}{2(P_{90} - P_{10})} \qquad (41)$$

فإذا كانت $\beta = 0.263$ فإن المنحنى معتدل التفرطح ، أما إذا كانت $\beta < 0.263$ فإن المنحنى مدبب ، وإذا كانت $\beta > 0.263$ فإن المنحنى متفرطح .

مثال (45) : إذا علمت أن $P_{10} = 15.0909$ و $P_{90} = 78$ و $Q_1 = 23.3333$ و $Q_3 = 66$

أوجد معامل التفرطح المئوي .

الحـل :

من التعريف نجد أن

$$\beta = \frac{Q_3 - Q_1}{2(P_{90} - P_{10})} = \frac{66 - 23.3333}{2(78 - 15.0909)} = \frac{42.6667}{125.8182} = 0.3391$$

أي أن المنحنى التكراري لهذه البيانات متفرطح وذلك لأن 0.3391 أكبر من 0.263 .

1 - بفرض أن البيانات الآتية تمثل الدخل الأسبوعي لمائة موظف يعملون بشركة الفنادق :

62 69 67 68 59 56 61 57 77 62 75 63 55 64 60
65 72 65 61 68 73 65 62 75 80 60 57 61 57 67
76 65 58 65 64 66 61 69 76 72 57 75 68 81 64
68 71 72 58 73 55 73 79 81 56 69 64 66 65 65
74 66 68 73 65 65 60 65 80 66 80 68 55 66 71
73 74 68 59 69 55 67 65 67 63 72 73 73 75 75
67 56 67 62 65 75 62 63 63 59

والمطلوب :

أ- تكوين جدول توزيع تكراري يبين توزيع الدخل لهذه المجموعة .

ب- نسبة الدخل الأسبوعي لكل فئة من الفئات .

جـ- تمثيل البيانات بيانياً باستخدام :

1- المدرج التكراري
2- المضلع التكراري
3- المنحنى التكراري
4 - المنحنى المتجمع الصاعد والهابط .

2 - إذا كان الجدول الآتي يبين توزيع عدد الساعات الإضافية التي عملها مجموعة من الموظفين خلال شهر معين :

	10 -	12 -	14 -	16 -	18 - 20
عدد الموظفين	6	8	10	8	6

والمطلوب

أ- متوسط عدد الساعات التي عملها الموظفين خلال ذلك الشهر .

ب- وسيط عدد الساعات التي عملها الموظفين خلال ذلك الشهر .

جـ- الانحراف المعياري لعدد الساعات التي عملها الموظفين خلال ذلك الشهر .

د- ما هو شكل منحنى هذه البيانات ؟ ولماذا ؟

هـ - العزم الثاني والثالث حول المتوسط الحسابي . ماذا تستنتج ؟

3 - بفرض أن البيانات الآتية تظهر نتائج تقرير الأداء الوظيفي للعاملين بأحد المصانع بنهاية السنة الماضية :

الأداء الوظيفي	ممتاز	جيد جداً	جيد	متوسط	دون المتوسط
العدد	350	400	180	100	50

فإذا قررت إدارة المصنع منح مكافئات لخمسين موظفاً من الذين كان أدائهم الوظيفي ممتاز بصرف النظر عن الموقع والدرجة الوظيفية ، وضح كيفية اختيار هؤلاء الأشخاص ومن هم بالتحديد ، ثم أوجد الوسيط والمنوال للأداء الوظيفي .

4 - إذا علمت أن :

أ - المدى لمجموعة من القيم يساوي صفراً ، فما قيمة الانحراف المعياري ولماذا ؟

ب - الانحراف الربيعي لمجموعة من القيم يساوي صفراً ، فما قيمة المدى ؟

جـ - $\sum x_i^2 = 0$ لعدد من القيم، فأوجد الانحراف المعياري ومعامل الاختلاف لهذه البيانات .

د - وسيط درجات طالب في 6 اختبارات هو 72 ، وعلمت أن الانحراف المعياري لهذه الإختبارات هو صفر ، فما هو متوسط درجاته في هذه المواد ، ولماذا ؟

5 - مثل البيانات التالية بيانياً بالرسم المناسب ، وأيهما أكثر تجانساً ولماذا ؟

الفترات	1 - 3	4 - 6	7 - 9	10 - 12	13 - 15
التكرار (أ)	5	8	10	4	2
التكرار (ب)	8	12	6	10	3

6 - إذا كان المتوسط الحسابي والانحراف المعياري لمجموعتين كالآتي :

البيان	المتوسط الحسابي	الانحراف المعياري	العدد
المجموعة الأولى	18	2	50
المجموعة الثانية	8	4	35

فاحسب الانحراف المعياري للمجموعتين معاً .

7 - من البيانات الآتية :

2 7 14 16 6 15 18 11 10 9 11 18 8 6 14 12 8

6 8 17 12 5 7 4 6 9 10 12 2 3 11 7 20 16 19

13 12 4 11 18 13 8 11 6 11 17 14 12 15 8 .

أ - كون جدول توزيع تكراري ذي فترات متصلة على أن تكون بداية الفترة الأولى تساوي 2 .

ب - مثل جدول التوزيع التكراري بيانياً باستخدام المدرج التكراري .

جـ - أوجد الانحراف الربيعي ومعامل الالتواء الثاني .

د - أوجد التكرار النسبي والمئوي .

8 - الجدول التالي يبين الدخل اليومي لثلاثين منتجاً بأحد المصانع .

الدخل	4-2	6-4	8-6	10-8	12-10
التكرار (f)	4	6	10	6	4

والطلوب إيجاد :

أ - الوسيط . ب - معامل الاختلاف .

جـ- معامل الالتواء الأول وما هو شكل المنحنى ولماذا ؟

9 - الجدول التالي يبن عدد القطع المصنوعة خلال أسبوع في مصنع معين :

76 77 75 75 83 70 74 69 79 71 79 77 74 75 76 71 73

73 79 75 80 78 74 71 70 81 77 67 71 81 75 77 74 75 77

75 74 75 77 68

أ - ضع البيانات في جدول تكراري مناسب .

ب - أوجد الوسط الحسابي والوسيط من البيانات الأولية ثم من الجدول التكراري وقارن بين الإجابتين .

10 - البيانات التالية تمثل المعدلات التراكمية لثلاثين طالباً :

2.2 2.0 1.9 1.8 2.1 2.6 2.0 1.6 1.5 2.6 1.4 2.0 1.4

1.6 2.1 2.3 2.0 2.2 2.4 2.2 2.0 1.9 2.5 2.9 2.4 2.5

. 2.2 2.3 1.7 2.2

أ - كون جدول تكراري مستخدماً 8 فترات (الفترات 1.55 - 1.35 ، 1.75 - 1.55 ، ... وهكذا) .

ب - كون جدول التوزيع التكراري التراكمي والنسبي .

جـ - مثل الجدول بيانياً باستخدام :

1 - المدرج التكراري . 2 - المضلع التكراري . 3 - المضلع التكراري التراكمي.

11 - البيانات التالية تمثل المدة (بالأشهر) التي يستغرقها خمسون مريضاً للشفاء من مرض معين :

6.1 5.7 4.9 3.1 4.5 2.9 2.7 3.8 5.1 2.5 3.6 4.3 5.6

4.4 2.5 5.6 5.1 3.7 4.2 4.9 3.5 2.1 4.0 6.2 1.8 3.6

3.6 4.1 4.0 3.7 2.9 2.2 4.8 3.9 4.6 3.1 2.8 7.3 3.5

. 2.5 4.9 3.7 4.2 2.8 3.4 1.6 3.9 4.0 3.9

والمطلوب وضع هذه البيانات في جدول توزيع تكراري وتمثيلها بيانياً بالرسم المناسب .

12 - إذا كانت البيانات التالية تمثل الأنفاق الأسبوعي لثمانية أشخاص على السلع الضرورية بالدينار :

46 , 52 , 47 , 45 , 48 , 56 , 52 , 50

أ - أوجد :

1 - المتوسط الحسابي . 2 - الوسيط . 3 - المنوال (إن وجد) .

4 - المدى . 5 - الانحراف المعياري . 6 - معامل الاختلاف .

7 - معامل الالتواء (α_3) .

ب - كرر الفقرة (أ) في الحالات التالية :

1 - إذا انخفض استهلاك كل شخص بمقدار 5 دنانير . ماذا تستنتج ؟

2 - إذا زاد استهلاك كل شخص بمقدار 5 دنانير . ماذا تستنتج ؟

جـ - أوجد المتوسط الهندسي والمتوسط التوافقي .

13 - إذا كان عدد أيام الغياب لمجموعة من الموظفين كما يلي :

عدد أيام الغياب (x)	2	3	4	5	المجموع
عدد الموظفين (f)	1	4	3	2	10

أوجد :

أ - المتوسط الحسابي ب - الوسيط جـ - المنوال د - الانحراف المعياري

هـ - معامل الالتواء و - المتوسط الهندسي ز - المتوسط التوافقي .

14 - الجدول التالي يوضح المصروفات الشهرية بالألف دينار لعدد من المنشآت :

المصروفات	1 -	5 -	9 -	13 -	17 -	21 -	25- 29	المجموع
عدد المنشآت	4	6	10	14	10	6	4	30

أوجد :

أ - المتوسط الحسابي ب - الوسيط جـ - الربيع الأدنى د - الربيع الأعلى

هـ - معامل التفرطح و- معامل الالتواء ز - معامل الاختلاف .

ثم مثل البيانات بيانياً باستخدام المنحنى التكراري والمنحنى التكراري المتجمع الصاعد النسبي والمئوي .

15 - إذا علم أن الوسيط والمنوال يساوى 7 و 6.7 على الترتيب لجدول التوزيع التكراري التالي :

الفترات	4 - 2	6 - 4	8 - 6	10 - 8	12 - 10
التكرار	4	f_1	10	f_2	5

المطلوب :

أ - إيجاد قيمة f_1 ، f_2 . ب - معامل الاختلاف . جـ - معامل الالتواء .

16 - أكمل الجدول التالي :

الفترات	التكرار	التكرار النسبي	التكرار المتجمع الصاعد
6 - 4	–	–	3
- 6	–	0.125	–
- 8	–	–	18
- 10	–	0.25	–
- 12	9	–	–
- 14	–	–	–
18 - 16	–	0.0625	48

ثم أوجد كلاً من :

أ - معامل الاختلاف ب - معامل الالتواء جـ - الربيع الأعلى والمئوي الخامس والسبعون د - العشير الثاني والمئوي العشرون هـ - ما هي أهم خصائص هذا التوزيع ؟ ولماذا ؟ و - مثل البيانات بيانياً بالرسم المناسب .

17 - الجدول التالي يبين توزيع درجات 40 طالب :

الدرجة	5 -	15 -	25 -	45-35
عدد الطلبة	6	20	10	4

أوجد :

أ - المتوسط الحسابي .

ب - الوسيط .

جـ - الانحراف المعياري .

د - معامل الالتواء ثم بين نوعه .

هـ - المتوسط الهندسي والمتوسط التوافقي .

18 - أكمل الجدول التالي :

الفترات	التكرار f_i	التكرار النسبي	التكرار التراكمي (المتجمع)
4 - 2	4	-	-
6 - 4	-	-	20
8 - 6	-	-	-
10 - 8	-	0.20	-
12 - 10	6	0.15	-
المجموع	-	-	

ثم أوجد :

أ - الوسيط (Median) ب - معامل الاختلاف (c . v) .

19 - المعطيات التالية تمثل تركيز الهيموجلوبين في الدم لخمسة وعشرين مريضاً من المرضى الذين يترددون على العيادة المجمعة بزاوية الدهماني .

13.2 13.0 6.5 7.4 9.7 9.1 15.1 12.9 6.0 11.9 10.1
17.9 14.6 13.2 13.6 15.8 16.0 13.4 11.2 10.5 9.5 8.9
10.9 12.1 12.8

أ - وضح المعالم الأساسية لهذه المعطيات وذلك من خلال وضعها في جدول توزيع تكراري يتكون من 5 فترات متساوية الطول .

ب - مثل المعطيات بالرسم المناسب .

جـ - كون جدول التوزيع التكراري النسبي .

د - إذا علمت أنه إذا قل تركيز الهيموجلوبين عن 10 فيكون الشخص مصاباً بفقر دم حاد ، فأوجد نسبة المصابين بهذا المرض .

هـ - كون جدول التوزيع التكراري المتجمع الصاعد ، ثم احسب التكرار المتجمع الصاعد النسبي

20 - الجدول الآتي يبين الزمن (بالأشهر) الذي أستغرقه مجموعة من الأطفال حتى السير لأول مرة بمدينة . أوجد المتوسط الحسابي والوسيط والمنوال . هل هذه النتائج تحقق العلاقة :

المتوسط الحسابي - المنوال = 3(المتوسط الحسابي - الوسيط)

الزمن بالأشهر	عدد الأطفال
- 7.5	2
- 8.5	6
- 9.5	10
- 10.5	28
- 11.5	50
- 12.5	56
- 13.5	23
- 14.5	8
16.5 - 15.5	3

21 - البيانات التالية تبين كميات حمض البول في دم 36 شخص في مستشفى الزهراء للكلى :

531 456 450 280 202 209 471 466 498 490 482 377

246 325 364 449 455 439 317 325 340 357 367 400

300 371 284 225 314 371 398 218 405 259 266 232

و المطلوب :

أ - وضع هذه البيانات في جدول توزيع تكراري ذي 7 فترات متساوية الطول .

ب - احسب المتوسط الحسابي و الوسيط و المنوال من البيانات الأولية .

جـ - احسب المتوسط الحسابي و الوسيط و المنوال من جدول التوزيع التكراري .

د - قارن بين الإجابات في الفقرتين ب ، جـ .

22 - أوجد الوسط الحسابي و الوسيط و المنوال و الانحراف المعياري للقراءات الآتية :

7 , 5- , 3 , 4 , 6- , 2 , 8 , 5-

23 - سحبت عينتان من مجتمعين فأعطتا النتائج الآتية :

العينة الأولى : $\sum_{i=1}^{10} x_i = 60 \quad , \quad \sum_{i=1}^{10} x_i^2 = 390$

العينة الثانية : $\sum_{i=1}^{8} y_i = 56 \quad , \quad \sum_{i=1}^{8} y_i^2 = 420$

أوجد المتوسط الحسابي و الانحراف المعياري لكل عينة ، أي العينتين أكثر تجانساً ؟ إذا دمجت العينتان ، ما المتوسط الحسابي للمجموعة الناتجة .

24 - القراءات الآتية هي درجات 10 طلبة في اختبار معين :

56 . 65 . 62 . 65 . 65 . 63 . 65 . 68 . 70 . 72

أوجد كلاً من المتوسط الحسابي و الوسيط و المنوال والتباين والانحراف المعياري و معامل الاختلاف لهذه البيانات .

25 - الجدول الآتي يبين أوزان عينة من 10 أطفال . أوجد كلاً من المتوسط الحسابي و الوسيط و المنوال و التباين و الانحراف المعياري :

الرقم :	1	2	3	4	5	6	7	8	9	10
الوزن :	13.2	15.4	13.0	16.6	16.9	13.6	14.4	15.0	14.6	13.1

26 - إذا كان المتوسط الحسابي والمدى للبيانات التالية : $b > 12 > 7 > a$ يساوي 10 و 11 على الترتيب فأوجد قيمة كل من a و b .

27 - إذا علمت أن الوسيط لثلاثة أعداد يساوي 30 ، والمتوسط الحسابي للعدد الأصغر والوسيط يساوي 25 ، والمتوسط الهندسي للعددين الأصغر والأكبر يساوي $20\sqrt{3}$. أوجد قيمة الأعداد الثلاثة .

الفصـل الثـاني

الاحتمـــالات

Probabilities

2 . 1 مقدمـــة Introduction

بالإضافة إلى التطبيقات العديدة لنظرية الاحتمالات يدخل مفهوم الاحتمال في حياتنا ومعاملاتنا اليومية ، فمثلاً غالبا ما نسمع ونقول التعبير التالي : " انه من المحتمل أن تمطر السماء غدا " أو " إنه من المرجح وصول الطائرة متأخرة هذا المساء " أو " إن الفرصة جيدة أمام الطالب للنجاح في مادة الإحصاء " أو " احتمال فوز فريق لكرة القدم على فريق آخر هو كذا " إن كل تعبير من التعبيرات السابقة مبني على مفهوم الاحتمال ، أي ترجيح حدوث حدث معين في المستقبل غير مؤكد الوقوع وبالتالي فبالاحتمال نعبر عن مقدار ثقتنا في وقوع هذا الحدث مستقبلاً.

فمصطلح " الاحتمال " يعني إذًا مقدار ثقتنا في إمكانية حدوث شئ غير مؤكد الوقوع . وبالرغم من أن مفهوم الاحتمال شائع بيننا وجزء طبيعي من معاملاتنا اليومية إلا انه لا يوجد تفسير علمي واحد متفق عليه ، ومقبول لمصطلح " الاحتمال " لدى جميع الإحصائيين والمختصين في هذا المجال . وخلال العقود الماضية كان كل تفسير لمفهوم (الاحتمال) من قبل بعض الإحصائيين يلاقي انتقاداً شديداً من الآخرين . وفي الواقع إن ما يعنيه مصطلح (الاحتمال) مازال موضع جدل وخلاف في كثير من المناقشات الفلسفية التي لها اتصال بأساسيات الإحصاء . وبالتالي سوف نتعرض إلى ثلاثة تفسيرات مختلفة لمفهوم الاحتمال ، حيث أن كل منها من الممكن أن يكون مفيداً عند تطبيق نظرية الاحتمال في المسائل العملية .

تفسير الاحتمال تكراراً نسبيا :

في كثير من المسائل العلمية يفسر احتمال حدوث نتيجة معينة على أنه التكرار النسبي لتلك النتيجة عندما تتكرر تجربتها عدداً كبيراً من المرات تحت ظروف متشابهة . وعلى ضوء هذا المفهوم إذا كان عدد الحالات التي يتحقق فيها الحدث A مثلاً هو m وكانت n هي العدد الكلي للحالات الممكنة لحدوث الحدث A فإن احتمال حدوث الحدث A الذي يرمز له بالرمز $P(A)$ هو

النسبة بين m و n وذلك عندما تكبر n (تقترب من ما لانهاية) ، أي أنه نهاية التكرار النسبي

بمعنى :

$$P(A) = \lim_{n \to \infty} \frac{m}{n}$$

ويسمى هذا التعريف أحياناً بالتعريف التجريبي لمصطلح " الاحتمال " . وبذلك يتضح جلياً أن قيمة الاحتمال تعتمد على n . فمثلاً ، احتمال الحصول على صورة عند إلقاء قطعة نقود مترنة يساوي $\frac{1}{2}$ والسبب في ذلك أن التكرار النسبي لعدد الصور عند إلقاء قطعة النقود عدداً كبيراً من المرات وتحت ظروف متشابهة يجب أن يكون $\frac{1}{2}$ تقريباً . وبعبارة أخرى أننا افترضنا بأن نسبة المرات التي نحصل فيها على صورة تساوي $\frac{1}{2}$ تقريباً . وفي الواقع تعتبر الشروط التي ذكرت في هذا المثال واستخدامها أساساً لتعريف مفهوم مصطلح " الاحتمال " غامضة إلى حد ما ، والسبب في ذلك يرجع إلى الاعتبارات التالية :

(أ) - يشترط أن تُلقى العملة عدداً كبيراً من المرات، ولكن في الواقع لا يوجد دليل محدد لعدد هذه المرات حتى يمكن اعتباره كبيراً بشكل كاف .

(ب) - عند القول بوجوب إلقاء العملة تحت ظروف متشابهة في كل مرة لا يوجد وصف دقيق لهذه الظروف. وفي الحقيقة أن الظروف التي ألقيت فيها العملة يجب ألا تكون متطابقة بالكامل في كل مرة والسبب يرجع إلى أن النتائج سوف تكون نفسها .وبالتالي من الممكن أن تكون جميعها صور أو جميعها كتابات . وفي الواقع من الممكن وجود شخص ماهر يمسك العملة بطريقة معينة ويلقيها تكراراً بحيث يحصل على صورة في كل مرة تقريباً ، وعليه أن عملية الإلقاء يجب أن لا تكون تحت تصرفنا بالكامل ولكن يجب أن تكون بطريقة عشوائية صرفه .

(جـ) - علاوة على ذلك عند القول بأن التكرار النسبي للصور يجب أن يكون مساوياً $\frac{1}{2}$ تقريباً لم يحدد بحد معين مسموح به للاختلاف عن $\frac{1}{2}$. فمثلاً عند إلقاء قطعة نقود 1000 مرة سنكون مندهشين عند الحصول على صورة في 500 مرة ، أي أننا لا نتوقع الحصول على هذا العدد ولكن نتوقع أن يكون قريباً منه (بالزيادة أو النقص) .وبالتالي يجب أن نكون قادرين على صياغة عبارة دقيقة ترجح مختلف الأعداد الممكنة للصور، ولكن هذا الترجيح من الضروري أن يكون معتمداً على المفهوم العميق الذي نود تعريفه لمصطلح الاحتمال .

(د) – أيضا من عيوب تفسير الاحتمال على أنه تكرار نسبي هو انطباقه على الأقل من حيث المبدأ على المسائل التي يمكن تكرارها عدداً كبيراً من المرات المتشابهة ، وفي الواقع هناك العديد من المسائل المهمة التي ليست من هذا النوع ، فمثلاً لا يمكن تطبيق هذا التفسير مباشرة على احتمال أن شخصاً معيناً سوف يتزوج خلال السنتين القادمتين .

التفسير التقليدي للاحتمال :

هذا التفسير مبنى على أساس مفهوم النتائج ذات الفرص المتساوية ، فمثلاً عند إلقاء قطعة نقود معدنية متزنة مرة واحدة يكون هناك نتيجتان ممكنتان إما صورة وإما كتابة . فإذا افترضنا أن لهاتين النتيجتين فرصاً متساوية في الحدوث ، فهذا يعني أن لهما احتمالاً متساوياً . وحيث إن مجموع الاحتمالات يجب أن يكون مساوياً للواحد الصحيح ، كما سنرى فيما بعد ، وعليه فكلا الاحتمالين للصورة والكتابة يساوي $\frac{1}{2}$. وبصفة عامة إذا كانت نتيجة تجربة ما ستكون واحدة من بين n من النتائج الممكنة لها وكانت لهذه النتائج فرص متساوية في الحدوث فإن احتمال حدوث كل نتيجة يساوي $\frac{1}{n}$. ولكن هناك مشكلتان أساسيتان عندما نحاول تفسير مصطلح الاحتمال من وجهة نظر الفرص المتساوية . المشكلة الأولى هي أن مفهوم النتائج بفرص متساوية جوهرياً مبنى على مفهوم الاحتمال الذي نحاول تعريفه ، وبذلك نكون قد عرفنا الاحتمال بدلالة الاحتمال . لأن القول بأن النتيجتين الممكنتين لهما فرصة متساوية في الحدوث هو نفس القول بأن النتيجتين لهما نفس الاحتمال . والمشكلة الثانية هي عدم وجود طريقة منتظمة في إعطاء الاحتمالات للنتائج التي نفترض أن ليس لها فرصاً متساوية . وخاصة في العلوم الطبيعية ، فمثلاً لا يمكن الاعتماد على فكرة مبدأ تكافؤ الفرص لتحديد نوع ، مولود ذكر أو أنثى. بينما عند إلقاء قطعة نقود أو مكعب نرد فإن النتائج الممكنة لهذا النوع من التجارب يمكن النظر إليها على أن لها فرصاً متساوية ، وذلك لأن طبيعة التجربة هكذا .

التفسير الذاتي للاحتمال :

ينطوي هذا التفسير على أن الشخص الذي يعطي احتمالاً معيناً لنتيجة ما في تجربة معينة ، يمثل وجهة نظره من حيث ترجيح حدوث تلك النتيجة . إن مثل هذا الحكم سوف يكون مبنياً على اعتقاد ومعلومات الشخص عن ظروف تلك التجربة ،حيث من الممكن وجود شخص آخر له اعتقاد مختلف أو معلومات مختلفة عن ظروف تلك التجربة . وبالتالي سوف

عن الاحتمال الشخصي لأشخاص معينين بدلاً من التحدث عنه كاحتمال حقيقياً للنتيجة .لذلك فهذا النوع من الاحتمالات يعبر عن قوة عقيدة (أو حدس) شخص ما ، اتجاه ظاهرة معينة ، وبالتالي فإن قيمته تختلف من شخص إلى آخر. فالاحتمال الذاتي لا يعتمد على أساس رياضي ويختلف من شخص لآخر ولكن غالباً ما يكون مبنياً على الخبرة أو التجربة . فمثلاً عند إلقاء قطعة نقود مرة واحدة ، فإن الشخص الذي ليست له معلومات شخصية عن قطعة النقود ، والطريقة التي ألقيت بها من الممكن أن يعتبر أن للصورة والكتابة فرصاً متساوية في الظهور ، وعليه فإن احتماله الشخصي سيكون $\frac{1}{2}$ للحصول على صورة أو كتابة . ولكن الشخص الذي قام بالعملية بنفسه من الممكن أن يكون له شعور بأن فرصة الحصول على صورة أكثر من فرصة الحصول على كتابة ، ولكي يكون لهذا الشخص قدرة على إعطاء احتمالات شخصية للنتائج يجب أن يعبر عن شعوره بلغة الأعداد .فإذا افترضنا على سبيل المثال انه يعتقد أن فرصة الحصول على صورة مثل فرصة الحصول على بطاقة حمراء عند سحبها من صندوق به أربع بطاقات حمراء وبطاقة زرقاء ، وحيث انه سوف يعطى احتمال $\frac{4}{5}$ لإمكانية الحصول على بطاقة حمراء فيجب عليه أيضاً إعطاء الاحتمال $\frac{4}{5}$ لإمكانية الحصول على صورة .

ولكن هناك مشكلتان رئيسيتان للتفسير الشخصي للاحتمال ، أولاهما هي أن الحكم الشخصي لترجيح عدد لانهائي من الأحداث يجب أن يكون متناسقاً بالكامل وخالياً من التناقض، ولكن يبدو انه أمر غير ممكن من الناحية البشرية . وثانيهما أن التفسير الشخصي لا يعطي أساساً موضوعيا لشخصين يعملان معاً للوصول إلى تقويم مشترك في مجال ذي اهتمام مشترك . إن النظرية الرياضية للاحصاء والاحتمال تلعب دوراً مهماً في الاختبارات والقرارات وسوف نعرضها في هذا الكتاب ، بغض النظر عن الجدل المحيط بهذه التفسيرات المختلفة لمفهوم مصطلح الاحتمال . هذه النظرية صحيحة ومفيدة من الناحية التطبيقية ، بصرف النظر عن تفسير مصطلح الاحتمال الذي سيستخدم في مسألة معينة . إن النظريات والأساليب التي سوف نعرضها في هذا الكتاب ستكون ذات أهمية وفائدة في جميع الظواهر تقريباً . خصوصاً المتعلقة بتصميم وتحليل عملية التجربة .

2 - 2 التجارب العشوائية Random Experiments

إن نظرية الاحتمال تختص بالنتائج أو الأحداث المختلفة والممكنة الحدوث التي يمكن الحصول عليها عند إجراء تجربة معينة ، وان تعبير " التجربة " المستخدم في نظرية الاحتمال هو وصف فعلي لأي عملية تكون نتيجتها غير معروفة مسبقاً بشكل أكيد . ومن الأمثلة على ذلك ما يلي :

1- إلقاء قطعة نقود تعتبر تجربة عشوائية لأننا نعلم أن لها نتيجتين إما صورة أو كتابة ولكن لا نستطيع تحديد نتيجة التجربة مسبقاً .

2 - رمي مكعب نرد مرة واحدة تعتبر تجربة عشوائية لأننا نعلم أن لها ستة نتائج ممكنة لكن لا نستطيع تحديد نتيجة التجربة مسبقاً .

3 - في تجربة إلقاء قطعة نقود 10 مرات من الممكن تحديد احتمال الحصول على أربعة صور على الأقل .

4 - في تجربة فحص صندوق به 100 جهاز كهربائي تم اختياره من بين مجموعة من الصناديق التي بها أجهزة متشابهة يمكن تحديد احتمال عدم وجود اكثر من جهاز به عطب في ذلك الصندوق .

5 -- في تجربة مراقبة حرارة الجو في موقع معين عند الساعة 12 ظهراً خلال مدة 90 يوماً متتالية يمكن تحديد الاحتمال بأن متوسط درجة الحرارة خلال تلك المدة سوف تكون أقل من قيمة معينة .

يمكن أن نلحظ من خلال هذه الأمثلة ، أن النتائج الممكنة للتجربة من الممكن أن تكون عشوائية أو غير عشوائية ، وفقاً للمعنى المألوف لهذه التعبيرات . إن السمات المهمة للتجربة هي أنه كل النتائج الممكنة يمكن تحديدها قبل أجراء التجربة ، وإعطاء الاحتمالات للمجموعات المختلفة للنتائج ذات الاهتمام . وكما سبقت الإشارة إلى وجود جدل في المعنى والتفسير المناسب لبعض الاحتمالات المعطاة لنتائج العديد من التجارب العشوائية . ولكن بمجرد إعطاء الاحتمالات لنتائج التجربة ، فإنه يوجد اتفاق كامل بين جميع الخبراء بأن النظرية الرياضية للاحتمال تعطى الطريقة المناسبة للدراسة المستقبلية لهذه الاحتمالات ، إن اغلب العمل تقريباً في النظرية الرياضية للاحتمال ، من الكتب المبدئية إلى البحوث المتقدمة له علاقة بالمسألتين الآتيتين :

i) طرائق تحديد الاحتمالات لأحداث معينة لكل نتيجة ممكنة في التجربة .

ii) طرائق تعديل احتمالات الأحداث عندما تتوفر معلومات إضافية ذات علاقة .

إن هذه الطرائق مبنية على أساليب رياضية معروفة ، وعلى ضوء ذلك فإننا في بعض الفصول من هذا الكتاب نعرض من خلالها هذه الأساليب والتي مع بعضها تكون قاعدة أساسية للنظرية الرياضية للاحتمال . ولكن قبل أن نخوض في الاحتمالات سوف نعرض وبإيجاز نظرية المجموعات التي لها صلة وثيقة بموضوع الاحتمالات .

2 - 3 نظرية المجموعات Set Theory

تُعرف المجموعة على أنها أي تجمع من الأشياء التي تشترك في صفة معينة ، وقد تكون هذه الأشياء كميات أو أعداداً أو أي شيء آخر معرفاً تعريفاً واضحاً . وعادة ما يرمز للمجموعة بأحد الحروف الهجائية الكبيرة مثل A أو B أو C أو . . . الخ . ويسمى كل عضو من أعضاء المجموعة عنصراً ويرمز له غالباً بأحد الحروف الصغيرة مثل a أو b أو c أو . . . الخ . حيث يتم حصر عناصر المجموعة بقوسين من النوع { } . والجدير بالملاحظة هنا أن ترتيب العناصر داخل المجموعة لا يؤثر على تعريفها ، ويرمز لانتماء عنصر ما إلى مجموعة معينة بالرمز $\in$ وعدم انتمائه بالرمز $\notin$. فمثلاً إذا كانت $A = \{ 1,2,3,4,5,6 \}$ و $B = \{ a,b,c,d \}$ فإننا نقول $5 \in A$ ، كما أن e لا تنتمي إلى B وتكتب $e \notin B$. ويمكن أن نعرف المجموعة بذكر الخواص التي تحقق عناصرها ، فمثلاً إذا كانت C هي مجموعة الأعداد الفردية فإن المجموعة C يمكن التعبير عنها كالآتي :

$C = \{ x :$ عدد فردي $x \}$ ونقرأ النقطتين : بعد الحرف x داخل القوسين " حيث أن " . ويستعمل مثل هذا التعريف عادة عندما تكون عناصر المجموعة لانهائية .

تعريف (1) : المجموعة الشاملة Universal Set

هي المجموعة التي تحتوي على جميع العناصر المكونة للظاهرة قيد الدراسة وبذلك فهي تختلف وتتغير حسب مجال البحث ، ويرمز لها بالرمز S . فمثلاً إذا كانت المجموعة الشاملة S هي مجموعة الأعداد الحقيقية فإن $S = \{ x :$ عدد حقيقي $x \}$.

تعريف (2) : المجموعة الخالية Empty Set

هي المجموعة التي لا تحتوي على عناصر ، بمعنى لا ينتمي إليها أي عنصر ، ويرمز لها بالرمز { } أو $\varnothing$ ، ومثالاً لذلك مجموعة أفراد المجتمع الذين يقطنون الشمس . ويلاحظ أن المجموعة الخالية $\varnothing$ موجودة في أي مجموعة أخرى . فعلى سبيل المثال $\varnothing$ موجودة في S .

تعريف (3) : المجموعة الجزئية Subset

هي المجموعة التي تكون جميع عناصرها موجودة في مجموعة أخرى . فإذا كانت $A = \{1, 2, 3, 4, 5, 6\}$ و $B = \{4, 5, 6\}$ فإن B مجموعة جزئية من A وتكتب $B \subset A$. لاحظ أن $A \subset A$ و $A \subset S$ و $\varnothing \subset A$. وبصفة عامة إذا كان عدد عناصر المجموعة هو n فإن عدد المجموعات الجزئية لهذه المجموعة هو 2^n . فمثلاً إذا كانت $A = \{a, b, c\}$ فإن عدد المجموعات الجزئية للمجموعة A يساوي $2^3 = 8$ وهذه المجموعات الجزئية هي :

$\{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\}, \{a, b, c\}, \{\ \}$

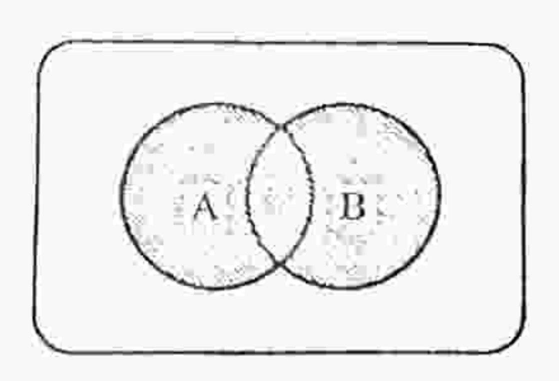
تعريف (4) : المجموعة المكملة Complement Set

إذا كانت A مجموعة ما ، فإن المجموعة التي تحتوى على جميع العناصر الموجودة في المجموعة الشاملة S وغير موجودة في المجموعة A تسمى المجموعة المكملة للمجموعة A ويرمز لها بالرمز A' أو A^c ، حيث $A' = S - A$ وهي تمثل مجموعة جميع العناصر التي في S وليست في A .

عمليات نظرية المجموعات Operation Set Theory

1 - الاتحاد The Union

إذا كانت A و B مجموعتين جزئيتين من S فإن المجموعة التي تتضمن جميع العناصر (النقاط) التي في A أو في B أو كليهما تعرف على أنها اتحاد A و B ويرمز لها بالرمز $A \cup B$ كما في شكل (1) أدناه :



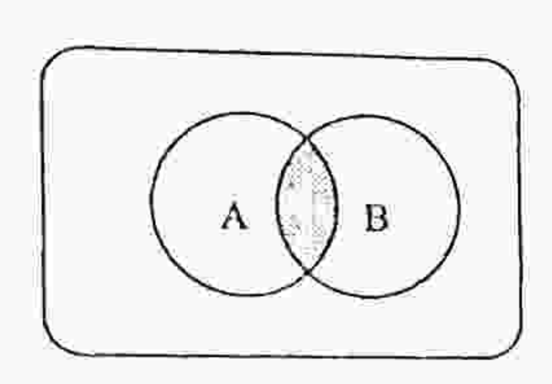
شكل (1) : $A \cup B$

وبصفة عامة إذا كانت $A_1, A_2, \dots, A_n$ تمثل مجموعات جزئيـة من S فـإن اتحـاد هذه المجموعات هي المجموعة التي تحتوى علـى جميـع العنـاصر (النقـاط) التـي تنتمـي علـى الأقل لواحدة من هذه المجموعات ويرمز لذلك بالرمز $\bigcup_{i=1}^{n} A_i$ حيث أن :

$$\bigcup_{i=1}^{n} A_i = A_1 \bigcup A_2 \bigcup \dots \bigcup A_n$$

2 - التقاطع The Intersection

إذا كانت A و B مجموعتين جزئيتين فـي S فـإن المجموعـة التـي تتضمـن العنـاصر المؤلفة من A و B تعرف على أنها تقاطع A مـع B ويرمـز لذلـك بـالرمز $A \cap B$ أو ببسـاطة AB كما في شكل (2) .



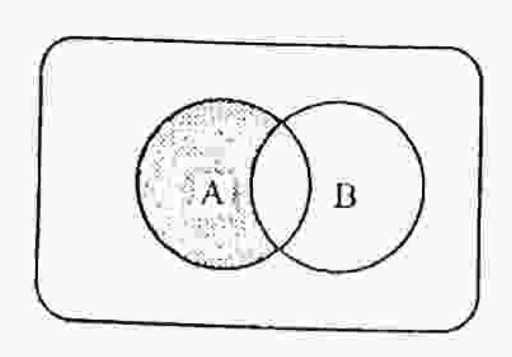
شكل (2) : $A \cap B$

وبصورة عامة إذا كانت $A_1, A_2, \dots, A_n$ مجموعات جزئية من S فـإن تقـاطع هذه المجموعات هي المجموعة المؤلفة من العناصر (النقاط) المشـتركة فـي جميـع هذه المجموعات ويرمز لذلك بالرمز $\bigcap_{i=1}^{n} A_i$ حيث أن $\bigcap_{i=1}^{n} A_i = A_1 \bigcap A_2 \bigcap \dots \bigcap A_n$.

3 - الفرق The Difference

إذا كانت A و B مجموعتين جزئيتين في S فإن الفرق بينهما ، والذي يرمز لـه بـالرمز A - B ، هو المجموعة التي تتضمن جميع النقاط الموجودة في A إلا أنهـا غيـر موجـودة فـي B

ويمكن التعبير عن ذلك بالمجموعات كما يلي : $A - B = \{x : x \in A , x \notin B\}$ كمـا فـي الشكل (3) أدناه :



شكل (3) : **A - B**

مثال (1) : إذا كانت $S = \{(x,y) : 0 \le x \le 1 , 0 \le y \le 1\}$ وبفرض أن :

$$A = \{(x,y) : 0 \le x \le 1 , 0 \le y \le \frac{1}{2}\}$$

$$B = \{(x,y\ .0 \le x \le \frac{1}{2} , 0 \le y \le \frac{1}{2}\}$$

$$C = \{(x,y) : 0 \le x \le y \le 1\}$$

$$D = \{(x,y) : 0 \le x \le \frac{1}{2} , 0 \le y \le 1\}$$

وعليه فإن :

$$B \subset A \quad و \quad B \subset D \quad و \quad A \cap D = AD = B$$

$$A' = \{(x,y) : 0 \le x \le 1 , \frac{1}{2} < y \le 1\}$$

$$A - B = \{(x,y) : \frac{1}{2} < x \le 1 , 0 \le y \le \frac{1}{2}\}$$

$$B \cup C = D \cup C$$

إن العمليات التي تُجرى على المجموعات محكومة بقوانين وبديهيات تفسر العلاقات بين المجموعات ، فإذا كانت A و B و C ثلاث مجموعات جزئية من المجموعة الشاملة S فإن :

(1) قانون التبديل Commutative Law

ينص هذا القانون على ما يلي :

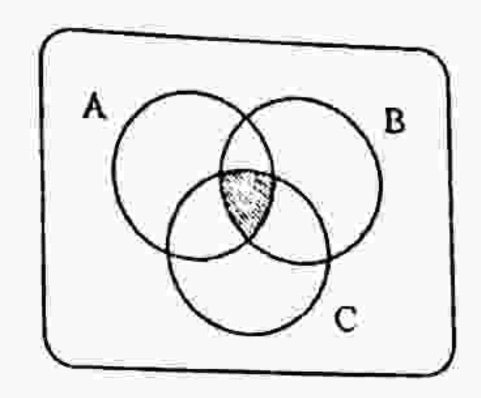
$$A \cap B = B \cap A \quad , \quad A \cup B = B \cup A$$

(2) قانون التنسيق Associative Law

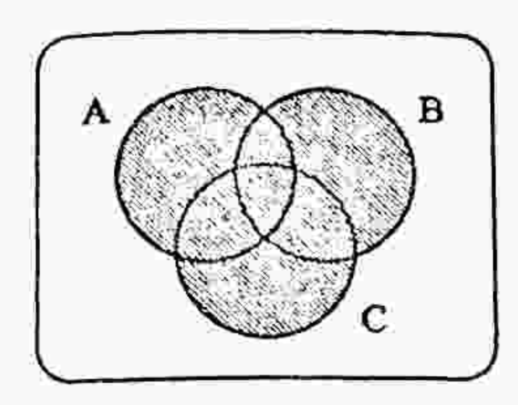
ينص هذا القانون على ما يلي :

ا - $(A \cup B) \cup C = A \cup (B \cup C)$

ب - $(A \cap B) \cap C = A \cap (B \cap C)$



شكل (4) : $(A \cap B) \cap C$

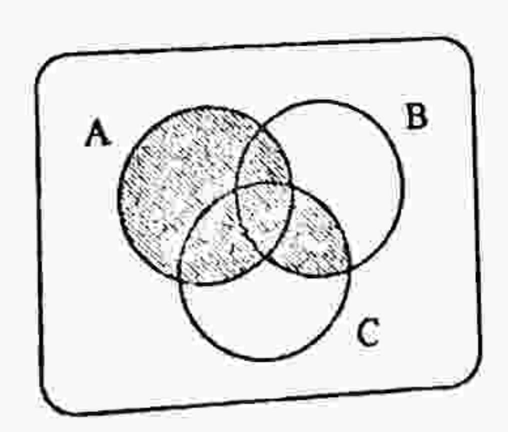


شكل (5) : $\mathbf{A \cup B \cup C}$

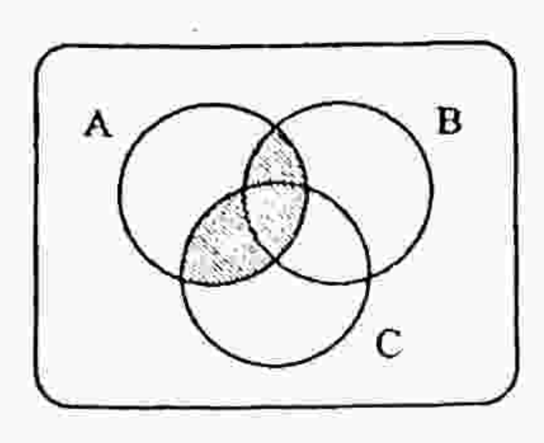
(3) قانون التوزيع Distributor Law

ينص هذا القانون على ما يلي : أ) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

ب) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$



شكل (6) : $A \cup (B \cap C)$



شكل (7) : $A \cap (B \cup C)$

(4) قانون المكملة Complementary Law

ينص هذا القانون على ما يلي :

أ) $A \cup A' = S \quad , \quad A \cap A' = \phi$

ب) $A \cup S = S$ و $A \cap S = A$ وذلك لان $A \subset S$

جـ) $(A')' = A$ و $A \cap \varnothing = \varnothing$ و $A \cup \varnothing = A$

(5) قانون الفرق Difference Law

أ) $A - B = A \cap B'$

ب) $A - B = A - (A \cap B) = (A \cup B) - B$

جـ) $(A \cup B) - C = (A - C) \cup (B - C)$

د) $(A \cap B) \cap (A - B) = \phi \quad , \quad (A \cap B) \cup (A - B) = A$

(6) قانون دى مورجان De Morgan's Law

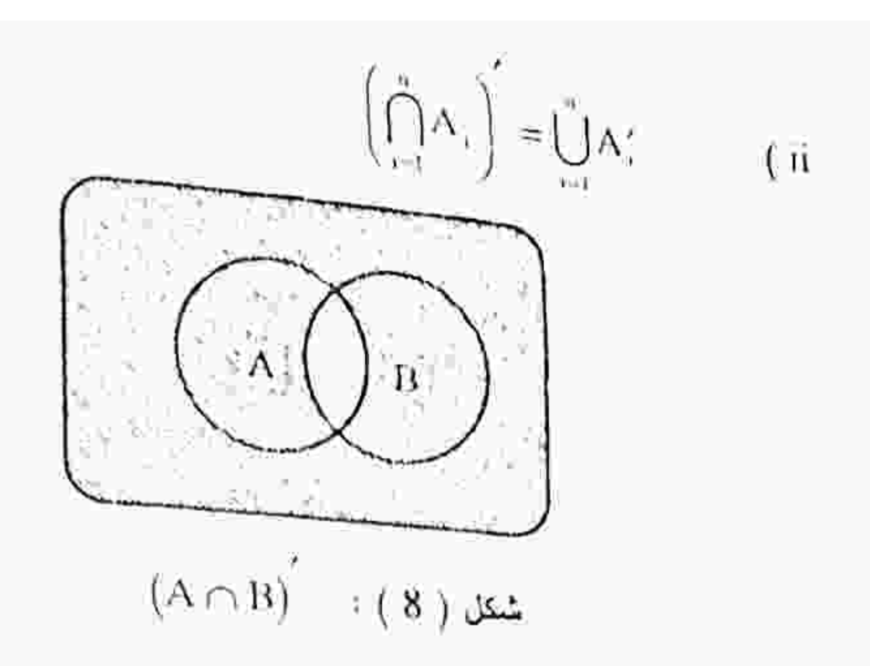
أ) $(A \cup B)' = A' \cap B'$

ب) $(A \cap B)' = A' \cup B'$

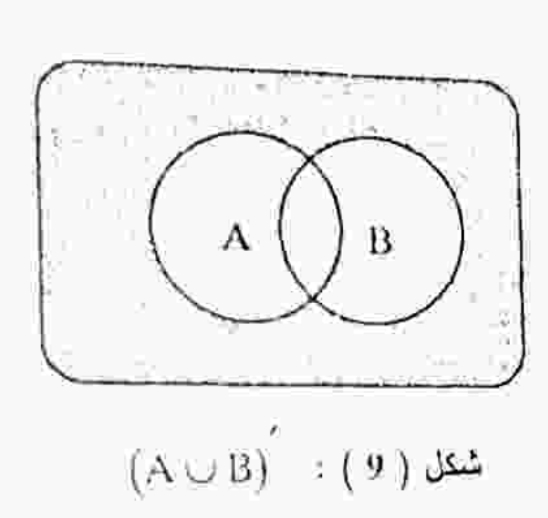
وبصفة عامة فإن :

i) $\left(\bigcup_{i=1}^{n} A_i \right)' = \bigcap_{i=1}^{n} A_i'$

$$\left(\bigcap_{i=1}^{n} A_i\right)' = \bigcup_{i=1}^{n} A_i' \qquad \text{(ii}$$



شكل (8) : $(A \cap B)'$



شكل (9) : $(A \cup B)'$

تعريف (5) : المجموعة القابلة للعد Countable Set

إذا أمكن عد أو ملاحظة عناصر مجموعة ما فإنها تكون قابلة للعد countable أما إذا تعذر ذلك تكون غير قابلة للعد uncountable . فمثلاً المجموعة $A = \{1, 2, 3, \cdots\}$ قابلة للعد بينما المجموعة $B = \{x : 0 < x < 10\}$ غير قابلة للعد .

تعريف (6) : المجموعة المحدودة (المنتهية) Finite Set

المجموعة المحدودة هي التي تحتوى على عدد معين من العناصر ، فمثلاً المجموعة $A = \{1, 2, \ldots, n\}$ حيث n عدد محدود تسمى مجموعة منتهية ، وأيضاً المجموعة $B = \{\ \}$ تعتبر مجموعة منتهية حيث أنها مجموعة خالية ، بينما المجموعة $B = \{x : 0 < x < 10\}$ والمجموعة $C = \{1, 2, 3, \cdots\}$ مجموعتان لانهائيتان .

2 . 4 فراغ (فضاء) العينة والأحداث Sample Space And Events

بعد أن عرضنا التفسيرات المختلفة للاحتمال وما يترتب عن كل تفسير عرضنا مراجعة بسيطة للمفاهيم الأساسية لنظرية المجموعات ، وفي هذا البند سوف نتعرض لتعريفيين مهمين هما فراغ العينة والحدث .

تعريف (7) : فراغ (فضاء) العينة Sample Space

فراغ العينة لتجربة عشوائية هو مجموعة جميع النتائج الممكنة لتلك التجربة ، ويرمز له بالرمز Ω (ويقرأ اوميقا) omega . ويقصد بالتجربة العشوائية هنا كل تجربة لا تكون نتيجتها معروفة مسبقاً بشكل حتمي .إن الأمر المهم في هذا التعريف هو احتواء Ω على جميع النتائج الممكنة للتجربة . أي أنها تقدم اكبر قدر ممكن من التفاصيل لهذه النتائج .

مثال (2) : إذا ألقيت قطعة نقود معدنية متزنة مرتين فإن فراغ العينة في هذه الحالة هو :

$$\Omega = \{HH, HT, TH, TT\}$$

حيث T ترمز للكتابة و H ترمز للصورة .

مثال (3) : إذا تم إلقاء مكعبي(زهرتي) نرد متزنين ومتمايزين مرة واحدة فإن :

$$\Omega = \{(i,j) : i = 1,2,3,4,5,6 \,;\, j = 1,2,3,4,5,6\}$$
$$= \{(1,1),(1,2),\ldots,(1,6),(2,1),(2,2),\ldots,(6,1),(6,2),\ldots,(6,6)\}$$

مثال (4) : إذا تم اختيار نقطة من الفترة $[0,1]$ ، فإن فراغ العينة في هذه الحالة هو : $\Omega = \{x : 0 \le x \le 1\}$ وهو يحتوى على عدد من النقاط غير القابلة للعد . ولكن إذا كانت التجربة تتضمن اختيار نقطة من مربع محدد بالنقاط $(0,0)$ ، $(1,0)$ ، $(1,1)$ ، $(0,1)$ فإن فراغ العينة : $\Omega = \{(x,y) : 0 \le x \le 1 \,,\, 0 \le y \le 1\}$ ، هو الآخر مجموعة غير قابلة للعد .

مثال (5) : إذا كانت التجربة تتضمن إلقاء قطعة نقود معدنية متزنة حتى ظهور صورة H في هذه الحالة يكون فراغ العينة هو : $\Omega = \{H, TH, TTH, TTTH, \ldots\}$ وهي مجموعة لا نهائية من النتائج وقابلة للعد .

نلحظ من الأمثلة السابقة أن كل نتيجة (outcome) ممكنة لتجربة عشوائية هي في الحقيقة عنصر ينتمي لفراغ العينة ويسمى نقطة فراغ العينة " A sample point " وأن لهذا الفراغ المزايا الآتية :

1 - إن فراغ العينة Ω مناظر للمجموعة الشاملة . وبمجرد اختياره يبقى ثابتاً وسيكون كل النقاش متعلقاً بهذا الفراغ .

2 - من الممكن أن تكون النقاط أو النتائج التي تحتويها Ω عددية أو وصفية .

3 - من الممكن أن يحتوى هذا الفراغ على عدد محدود أو غير محدود من العناصر .

تعريف (8) : فراغ العينة المنفصل (المتقطع) Discrete Sample Space

إذا احتوى فراغ العينة Ω على الأكثر على عدد من العناصر القابلة للعد يسمى فراغاً منفصلاً .

تعريف (9) : فراغ العينة المتصل Continuous Sample Space

إذا احتوى فراغ العينة Ω على مجموعة غير قابلة للعد سواء كانت منتهية أو غير منتهية يسمى فراغاً متصلاً .

ومن الأمثلة على فراغ العينة المنفصل الأمثلة رقم (2) و (3) و (5) سالفة الذكر بينما المثال رقم (4) يمثل فراغ عينة متصلاً . وعادة يمثل فراغ العينة المتصل تلك النتائج المقاسة على مقياس متصل ، مثل درجة الحرارة والزمن والسرعة والوزن والطول . . . الخ . بينما يمثل فراغ العينة المنفصل تلك النتائج للتجربة العشوائية التي يمكن التعبير عنها بأعداد صحيحة فقط مثل عدد الحوادث خلال فترة زمنية معينة .

تعريف (10) : الحدث The Event

الحدث هو مجموعة جزئية من فراغ العينة Ω . ويرمز للأحداث عادة بحروف كبيرة مثل A و B و C و . . . الخ . وإذا احتوى الحدث على عنصر واحد من عناصر فراغ العينة يسمى حدثاً بسيطاً (Simple Event) أما إذا احتوى على اكثر من عنصر واحد فيسمى حدثاً مركباً (Compound Event) . فمثلاً إذا كانت : $\Omega = \{ HH , HT , TH , TT \}$ و $A = \{ HH \}$ و $B = \{ HH , TT \}$ و $C = \{ HH , HT , TT \}$ فإن A يمثل حدثاً بسيطاً بينما الحدثان B و C مركبان .

تعريف (11) : الحدث المكمل Complement Event

الحدث المكمل للحدث A مثلاً هو الحدث الذي يحتوى على جميع نتائج التجربة العشوائية (عناصر فراغ العينة) التي لا يشملها الحدث الأصلي A ويرمز لذلك الحدث بالرمز A' أو A^c .

تعريف (12) : الحدث المؤكد Sure Event

هو الحدث الذي يحتوى على جميع عناصر فراغ العينة . فالحدث A مثلاً يكون حدثاً مؤكداً إذا كان $A = \Omega$.

تعريف (13) : الحدث المستحيل Impossible Event

هو الحدث الذي لا يحتوى على أي نتيجة من نتائج فراغ العينة . فالحدث A مثلاً يكون مستحيلاً إذا كان $A = \varnothing$.

تعريف (14) : الأحداث المتنافية Mutually Exclusive Events

الأحداث المتنافية هي الأحداث المانعة لبعضها البعض ، فيقال إن A و B حدثان متنافيان إذا كان حدوث أحدهما يمنع حدوث الآخر ، أي أنه من المستحيل وقوع الحدثين معاً بمعنى $A \cap B = \varnothing$.

وبصورة عامة تكون الأحداث $A_1, A_2, \ldots, A_n$ متنافية ثنائياً (مانعة لبعضها البعض) إذا كان $A_i \cap A_j = \phi$ لكل $i \neq j$ لجميع قيم i و $j = 1, 2, \ldots, n$.
إذن بلغة نظرية المجموعات فإن الأحداث المتنافية تعني المجموعات المنفصلة .

تعريف (15) : الأحداث المستقلة Independent Events

يقال أن الحدثين A و B مستقلان إذا كان حدوث أحدهما لا يؤثر ولا يتأثر بحدوث أو عدم حدوث الآخر . وبصورة عامة تكون الأحداث $A_1, A_2, \ldots, A_n$ مستقلة إذا كانت لا تؤثر ولا تتأثر ببعضها البعض .

مثال (6) : إذا تم اختيار مصباح كهربائي من بين مجموعة من المصابيح التي أنتجها مصنع معين وسجل الزمن الكلي بالساعات الذي عمره ذلك المصباح فإن أي عدد غير سالب سيكون نتيجة مناسبة لهذه التجربة ، وعليه يمكن كتابة فراغ عينة هذه التجربة بالصيغة التالية : $\Omega=\{\omega : \omega \geq 0\}$. فإذا كان الحدث A يمثل المصباح سيشتعل على الأقل 10 ساعات قبل أن يحترق فإن $A=\{\omega : 10<\omega<\infty\}$ ، حيث ω تمثل عمر المصباح بالساعات .

مثال (7) : إذا ألقيت قطعة نقود معدنية متزنة ومكعب (زهرة) نرد معاً مرة واحدة فإن فراغ العينة Ω يتضمن 12 عنصراً (نقطة) وهي :

$$\Omega=\{\omega_1,\omega_2,\cdots,\omega_{12}\}$$
$$=\{(H,1),(H,2),\cdots,(H,6),(T,1),(T,2),\cdots,(T,6)\}$$

وعليه إذا كان الحدث A يمثل صورة ورقماً زوجياً فإن :

$$A=\{\omega_2,\omega_4,\omega_6\}=\{(H,2),(H,4),(H,6)\}$$

وإذا كان الحدث B يمثل كتابة ورقماً اكبر من 3 فإن :

$$B=\{\omega_{10},\omega_{11},\omega_{12}\}=\{(T,4),(T,5),(T,6)\}$$

مثال (8) : بفرض أن التجربة تمثل عدد الوفيات نتيجة لحوادث الطرقات في مدينة طرابلس خلال شهر معين ، فإن أي عدد صحيح غير سالب سيكون نتيجة مناسبة لهذه التجربة ، وعليه فإن :

$$\Omega=\{0,1,2,3,\ldots\}$$

وإذا كان الحدث A يمثل عدد الوفيات الأقل من 150 فإن :

$$A=\{0,1,2,\cdots,149\}$$

2 . 5 مسلمات الاحتمال Axioms of Probability

إذا وجدت تجربة عشوائية بفراغ عينة Ω وكان A يمثل حدث معرف على هذا الفراغ فإن ما يتبادر إلى الذهن هو كيف يمكننا تقويم درجة عدم التأكد في A ، بمعنى كيف يمكننا حساب احتمال حدوث A ؟ لقد أشرنا إلى ما هو تفسير بسيط ويعتمد على الإدراك الحسي ومنها ما يعتمد على التجريد المبني على نظريات رياضية مختلفة ، وأوضحنا أن أحد هذه التفسيرات مبني على فكرة التكرار النسبي ، فإذا أعيدت التجربة n من المرات وكان m هو عدد

المرات التي ظهر فيها الحدث A فإنه سيتولد لدينا شعور بأن التكرار النسبي للحدث A أي $\frac{m}{n}$ سوف يستقر بالقرب من عدد معين وليكن P_n كلما اقتربت n من ∞ وبعبارة أخرى إننا نتوقع بأن $\lim_{n \to \infty} \frac{m}{n} = P_n$. إذن هناك مخاطرة إذا أخذنا بمفهوم التكرار النسبي تعريفاً لاحتمال الحدث A ، فكيف يمكننا التأكد من وجود نهاية لهذا التكرار ؟ أيضا ما قيمة n التي يجب أن تؤخذ قبل إعطاء الحدث A الاحتمال $\frac{m}{n}$ ؟ لذا وكما أشرنا سابقاً فإن احتمال حدوث الحدث A يجب أن لا يعتمد على الشخص القائم بالتجربة ولا على التكرار المعين m الذي تتم مشاهدته عند إعادة التجربة n من المرات . علاوة على ذلك وكما أشرنا في البداية في كثير من المسائل فإنه من غير العملي إعادة التجربة . وعليه نقول إن التعريف السابق لا يفي بموضوع دراسة الاحتمالات وإننا سنعتمد التعريف الرياضي للاحتمال الذي يعتمد على عدة فرضيات تتناسب مع فكرتنا الادراكية لمعنى الاحتمال .

تعريف (16) : إذا كانت Ω تمثل مجموعة النتائج الممكنة للتجربة العشوائية فإن الدالة P(.) تسمى احتمالاً إذا حققت الشروط الآتية :

أ) $0 \leq P(A) \leq 1$ لأي حدث A ينتمي إلى Ω .

ب) $P(\Omega) = 1$ و $P(\varnothing) = 0$

جـ) إذا كانت $\{A_n\}$ تمثل متوالية من الأحداث المتنافية $\{A_i \cap A_j = \phi , i \neq j\}$

فإن $P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$

هذه الشروط الثلاثة تسمى مسلمات الاحتمال التي من خلالها يمكن تحديد مفهوم الاحتمال . إن التعريف السابق للاحتمال هو تعريف رياضي ، ويدلنا على دوال المجموعة التي يمكن أن نطلق عليها تسمية دوال احتمال ولكنه لا يدلنا على قيمة دالة الاحتمال P(A) المعطاة للحدث A . وللحصول على قيم لاحتمالات الأحداث علينا أن نضع التجربة العشوائية في نموذج معين . إن كلمة العشوائية في الإحصاء تعني أننا نعطي الاحتمال لعناصر (نقاط) فراغ العينة بطريقة تتفق مع رغبتنا في معاملة جميع النتائج على أنها متساوية من حيث إمكانية حدوثها (مبدأ تكافؤ الفرص) .

مثال (9) : إذا ألقيت قطعة نقود معدنية مرة واحدة ، فإن $\Omega = \{H, T\}$ وكان الحدث A يمثل ظهور صورة فإن $A = \{H\}$ وعليه فإن $P(A) = \frac{1}{2}$.

مثال (10) : عند إلقاء قطعة نقود معدنية متزنة ثلاث مرات ، حدد عناصر فراغ العينة ثم أوجد احتمال حدوث الأحداث التالية :

أ - الحصول على صورة واحدة على الأقل .

ب - الحصول على ثلاثة وجوه متشابهة .

جـ - الحصول على صورتين على الأكثر .

الحل :

حيث أن قطعة النقود تم إلقائها ثلاث مرات فإن :

عدد عناصر فراغ العينة = $n(\Omega)$ حيث : $n(\Omega) = 2^3 = 8$

وهي : $\Omega = \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}$

أ - بفرض أن : حدث الحصول على صورة واحدة على الأقل = A

فإن العناصر التي تحقق الحدث A هي :

$$A = \{HHH, HHT, HTH, THH, HTT, THT, TTH\} \Rightarrow n(A) = 7$$

وعليه فإن

$$P(A) = \frac{n(A)}{n(\Omega)} = \frac{7}{8}$$

ب - بفرض أن : الحصول على ثلاثة وجوه متشابهة = B . فإن :

$$B = \{HHH, TTT\} \Rightarrow n(B) = 2$$

وبالتالي فإن :

$$P(B) = \frac{n(B)}{n(\Omega)} = \frac{2}{8}$$

جـ - بفرض أن : الحصول على صورتين على الأكثر = C فإن :

$$C = \{HHT, HTH, THH, HTT, THT, TTH, TTT\} \Rightarrow n(C) = 7$$

وعليه فإن :

$$P(C) = \frac{n(C)}{n(\Omega)} = \frac{7}{8}$$

مثال (11) : عند إلقاء مكعب (زهرة) نرد متزن مـرة واحدة حدد عناصر فراغ العينـة ثـم أوجد احتمال حدوث الأحداث التالية :

أ - الحصول على عدد زوجي . ب - الحصول على عدد اكبر من 2 واقل من أو يساوي 4

جـ - الحصول على عدد فردي . د - الحصول على عدد على الأقل يساوي 2 واقل من 6

الحل :

عدد عناصر فراغ العينة يساوي 6 وهى : $\Omega = \{1,2,3,4,5,6\}$

بفرض أن :

حدث الحصول على عدد زوجي = A

حدث الحصول على عدد اكبر من 2 واقل من أو يساوي 4 = B

حدث الحصول على عدد فردي = C

حدث الحصول على عدد على الأقل يساوي 2 واقل من 6 = D

وعليه فإن :

$$A = \{2,4,6\} \qquad \Rightarrow n(A) = 3$$
$$B = \{3,4\} \qquad \Rightarrow n(B) = 2$$
$$C = \{1,3,5\} \qquad \Rightarrow n(C) = 3$$
$$D = \{2,3,4,5\} \qquad \Rightarrow n(D) = 4$$

وعليه فإن :

أ - احتمال الحصول على عدد زوجي :

$$P(A) = \frac{n(A)}{n(\Omega)} = \frac{3}{6} = 0.5$$

ب - احتمال الحصول على عدد اكبر من 2 واقل من أو يساوي 4 :

$$P(B) = \frac{n(B)}{n(\Omega)} = \frac{2}{6}$$

جـ - احتمال الحصول على عدد فردي :

$$P(C) = \frac{n(C)}{n(\Omega)} = \frac{3}{6} = 0.5$$

د - احتمال الحصول على عدد على الأقل يساوي 2 واقل من 6 :

$$P(D) = \frac{n(D)}{n(\Omega)} = \frac{4}{6}$$

من خلال تعريف الاحتمال والمسلمات السابقة سوف نعرض العديد من الخواص للدالة $P(.)$ والتي سنعرضها في صورة نظريات .

نظريـــة (1) : لأي حدث A يكون : $P(A')=1-P(A)$

البرهان :

حيث أن : $A \cap A' = \phi$ و $A \cup A' = 1$

وعليه من المسلمتين (ب) و (جـ) نستنتج أن :

$$P(\Omega) = P(A \cup A') = P(A) + P(A')$$

ولكن $P(\Omega) = 1$ وعليه فإن $P(A') = 1 - P(A)$.

نظريـــة (2) : لأي حدثين A و B يكون :

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

البرهان :

من شكل (10) نلاحظ أن :

$$A \cup B = (AB') \cup (AB) \cup (A'B)$$

وحيث أن AB' ، AB ، $A'B$ أحداث متنافية فإنه من المسلمة (جـ) نستنتج أن :

$$P(A \cup B) = P(AB') + P(AB) + P(A'B)$$

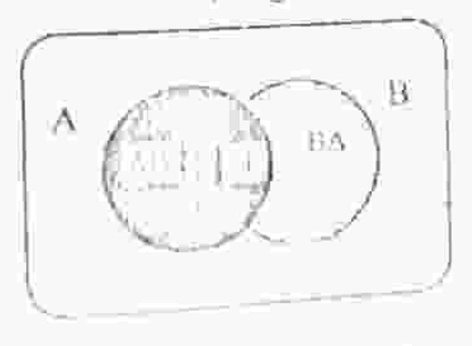
ولكن من شكل (10) أيضاً نلاحظ أن :

$$P(A) = P(AB') + P(AB)$$

و

$$P(B) = P(A'B) + P(AB)$$

وبالتعويض في $P(A \cup B)$ نحصل على النتيجة .



شكل (10) $A \cup B$

نتيجة (1) : إذا كان $A_1, A_2, \cdots, A_n$ تمثل متوالية من الأحداث فإن :

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i) - \sum_{i<j}\sum_{=2} P(A_i A_j) + \sum_{i<j<k}\sum\sum_{=2} P(A_i A_j A_k) + \cdots + (-1)^{n-1} P(A_1 A_2 A_3 \ldots A_n)$$

البرهان :

باستخدام الاستنتاج الرياضي الحظ أنه إذا كانت $n = 3$ فإن :

$$P(A_1 \cup A_2 \cup A_3) = P(A_1) + P(A_2) + P(A_3) - P(A_1 A_2) - P(A_1 A_3) - P(A_2 A_3) + P(A_1 A_2 A_3)$$

نظريـــة (3) : لأي حدثين A ، B يكون :

$$P(A \cup B) \leq P(A) + P(B)$$

البرهان :

حيث أنه من النظرية (2) :

$$P(A \cup B) = P(A) + P(B) - P(AB)$$

ومن المســلمة (أ) : $P(AB) \geq 0$

فإن $P(A \cup B) \leq P(A) + P(B)$

نتيجــة (2) : إذا كانت $A_1, A_2, \cdots, A_n$ تمثل أحداثاً فإن :

$$P\left(\bigcup_{i=1}^{n} A_i\right) \leq \sum_{i=1}^{n} P(A_i)$$

البرهان :

من النظرية (3) واستخدام الاستنتاج الرياضي يتم برهان هذه النتيجة .

نظريــة (4) : إذا كان $A \subseteq B$ فإن :

$$P(A) \leq P(B) \quad \text{و} \quad P(A'B) = P(B) - P(A)$$

البرهان :

حيث أنه إذا كانت $A \subseteq B$ فإن A و $A'B$ حدثان متنافيان واتحادهما هو B ، و [illegible]

من شكل (11) أدناه نجد أن :

$$P(B) = P[A \cup (A'B)] = P(A) + P(A'B)$$

$P(A'B) = P(B) - P(A)$

ومن ذلك نستنتج أن

ومن المسلمة (1) نجد أن $P(A'B) \geq 0$ وعليه فإن

$$P(B) - P(A) \geq 0 \quad \Rightarrow \quad P(A) \leq P(B)$$

شكل (11) : $A' \cap B$

نظريــة (5) :إذا كان A ، B حدثين فإن : $P(AB) \geq 1 - P(A') - P(B')$

البرهان :

من النظرية (1) نجد أن :

$$P(AB) = 1 - P((AB)')$$

$$= 1 - P(A' \cup B') \qquad ((A \cap B)' = A' \cup B' \text{ حيث أن })$$

$$\geq 1 - P(A') - P(B') \qquad (\text{من النظرية } (3))$$

لاحظ أنه يمكن تعميم هذه النظرية إلى n من الأحداث .

مثال (12) : إذا كان A ، B حدثان متنافيان وكـان $P(A) = 0.10$ ، $P(B) = 0.20$ فأوجد الاحتمالات التالية : $P(A')$ ، $P(B')$ ، $P(A \cup B)$ ، $P(A \cap B)$ ، $P(A' \cap B')$

الحل :

$$P(A') = 1 - P(A) = 1 - 0.10 = 0.90$$

$$P(B') = 1 - P(B) = 1 - 0.20 = 0.80$$

$$P(A \cup B) = P(A) + P(B) = 0.10 + 0.20 = 0.30$$

$$P(A \cap B) = 0$$

$$P(A' \cap B') = P\left((A \cup B)'\right) = 1 - P(A \cup B) = 1 - 0.30 = 0.70$$

مثال (13) : أخذت عينه عشوائية من طلبة قسم الإحصاء بكلية العلوم جامعة الفاتح حجمها 60 طالباً فوجد أن 15 طالباً مسجلين في مقرر ST305 و 12 طالباً مسجلين في مقرر ST401 و 6 طلبة مسجلين في المقررين معاً . فإذا تم اختيار أحد الطلبة عشوائياً من العينة المختارة فأوجد احتمال حدوث الأحداث التالية :

أ - أن يكون الطالب مسجل في أحد المقررين على الأقل .

ب - أن يكون الطالب غير مسجل في أياً من المقررين .

جـ - أن يكون الطالب مسجل في مقرر ST305 فقط .

الحل :

بفرض أن :

حدث اختيار طالب مسجل في مقرر A = ST305

حدث اختيار طالب مسجل في مقرر B = ST401

وعليه فإن :

$$P(A \cap B) = \frac{6}{60} \quad , \quad P(B) = \frac{12}{60} \quad , \quad P(A) = \frac{15}{60}$$

ا - احتمال أن يكون الطالب المختار مسجل في أحد المقررين على الأقل :

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = \frac{15}{60} + \frac{12}{60} - \frac{6}{60} = \frac{21}{60} = 0.35$$

ب - احتمال أن يكون الطالب المختار غير مسجل في أياً من المقررين :

$$P\left((A \cup B)'\right) = 1 - P(A \cup B) = 1 - 0.35 = 0.65$$

جـ - احتمال أن يكون الطالب المختار مسجل في مقرر ST305 فقط :

$$P(A \cap B') = P(A) - P(A \cap B) = \frac{15}{60} - \frac{6}{60} = \frac{9}{60}$$

مثال (14) : بفرض أن 80 % من الجرائم في مدينة معينة تحدث ليلاً و 90 % من هذه الجرائم تحدث في المنطقة الواقعة فيها تلك المدينة فماذا يمكن القول عن نسبة الجرائم التي تحدث ليلاً بتلك المنطقة ؟

الحل :

بفرض أن الحدث A يمثل الجريمة التي تحدث ليلاً ، والحدث B يمثل الجريمة حدثت بالمنطقة الواقعة فيها المدينة ، وعليه فإن :

$$P(A)=0.80 \quad , \quad P(B)=0.90$$
$$P(A')=0.20 \quad , \quad P(B')=0.10$$

وأن

$$P(A \cap B) \geq 1-0.20-0.10=0.70$$

وباستخدام النظرية (5) نجد أن :

إذن احتمال أن الجريمة تحدث ليلاً وبالمنطقة الواقعة فيها المدينة يساوي على الأقل 0.70

2 - 6 طرائق عد عناصر فراغ (فضاء) العينة : Sample Space Counting Methods

لقد عرفنا مما سبق أنه إذا كان فراغ العينة Ω مجموعة منتهية أي يحتوى على عدد محدود من النقاط وكانت هذه النقاط متساوية من حيث إمكانية حدوثها فإن احتمال حدوث الحدث A ($A \subseteq \Omega$) يساوى نسبة عدد نقاط ذلك الحدث إلى عدد نقاط فراغ العينة أي أن :

$$P(A)=\frac{n(A)}{n(\Omega)}$$

حيث $n(A)$ هو عدد النقاط في A و $n(\Omega)$ هو عدد النقاط في Ω . ولكن في كثير من التجارب العشوائية نجد أن عدد النتائج (نقاط أو عناصر فراغ العينة) في Ω سيكون كبيراً جداً وأن عملية وصفها (كتابة جميع النتائج الممكنة للتجربة العشوائية) غالباً ما يكون صعباً . في مثل تلك التجارب يبدو من المناسب إيجاد طرق لتحديد العدد الكلي للنتائج الممكنة بالفراغ Ω وللأحداث المختلفة فيها ، بدون أن يكون هناك ضرورة إلى وجود قائمة لجميع تلك النتائج . ولهذا سوف نتعرض في هذا البند إلى عدد من القواعد التي تلعب دوراً هاماً في تحديد تلك النتائج .

I - قاعدة الضرب Multiplication Rule

تنص هذه القاعدة على ما يلي : افترض لدينا تجربتين A و B للأولى n من النتائج المختلفة وللثانية m من النتائج المختلفة . عندئذ يكون عدد النتائج الممكنة للتجربتين هو معاً n m .

مثال (15) : لنفرض وجود ثلاثة طرق مختلفة تؤدى من المدينة A إلى المدينة B وخمسة طرق مختلفة تؤدى من المدينة B إلى المدينة C ، ما هو عدد الطرق التي تؤدى من المدينة A إلى المدينة C مروراً بالمدينة B ؟

الحل :

عدد الطرق التي تؤدى من المدينة A إلى المدينة C مروراً بالمدينة B يساوي :

$$n\,m = 3 \times 5 = 15$$

مثال (16) : كم عدد النقاط بفضاء العينة إذا تم إلقاء مكعبي (زهرتي) نرد متمايزين مرة واحدة ؟

الحل :

فإن عدد النقاط بفراغ العينة يساوي : $n\,m = 6 \times 6 = 36$

وذلك لأنه توجد ستة نتائج ممكنة لكل مكعب .

ويمكن تعميم قاعدة الضرب لتشمل k من التجارب كالآتي :

نتيجة (3) : أفترض لدينا k من التجارب $A_1, A_2, \ldots, A_k$ لهما $n_1, n_2, \ldots, n_k$ من النتائج المختلفة على التوالي . عندئذ يكون عدد النتائج الممكنة لهذه التجارب معاً هو :

$$\prod_{i=1}^{k} n_i = n_1\, n_2\, \ldots\, n_k$$

نتيجة (4) : إذا كانت التجربة تتضمن k محاولة حيث أن كل محاولة يمكن أن تحدث في واحدة من n من النتائج الممكنة فإن عدد النتائج الممكنة بفراغ العينة هو n^k .

إن هذه النتيجة مبنية على النتيجة (3) وذلك بوضع $n_1 = n_2 = \ldots = n_k = n$

مثال (17) : إذا كانت التجربة تتضمن إلقاء قطعة نقود معدنية متزنة 5 مرات فما هو عدد عناصر فراغ (فضاء) العينة ؟

الحل :

عدد النقاط بفراغ (بفضاء) عينة هذه التجربة بناء على النتيجة (4) هو :

$$n(\Omega) = n^k = 2^5 = 32$$

حيث : n = النتيجة الممكن الحصول عليها (صورة أو كتابة)

k = عدد مرات الإلقاء

مثال (18) : إذا كانت التجربة تتضمن عدد الهواتف التي يمكن تركيبها في مدينة قصر الأحيار علما بأن رقم الهاتف يتكون من أربعة أرقام على أن يكون أولها فردي فإن الرقم الأول له خمسة طرائق أما الأرقام الثلاثة الأخرى فيوجد عشرة طرائق لكل منها إذا كان التكرار مسموحاً به . وبالتالي فإنه من النتيجة (3) يكون عدد الهواتف التي يمكن تركيبها في مدينة قصر الأخيار مساو إلى :

$$5 \times 10 \times 10 \times 10 = 5000$$

II - قاعدة الجمع Addition Rule

تنص هذه القاعدة على ما يلي :

إذا كانت التجربة A تحدث في n من النتائج الممكنة والتجربة B في m من النتائج المحتملة وكانت التجربتان متنافيتين فإن عدد النتائج الممكنة من التجربة A أو التجربة B هو n + m نتيجة

نتيجة (5) : بصورة عامة إذا كانت التجربة A_1 تحدث في n_1 نتيجة والتجربة A_2 في n_2 نتيجة . . . والتجربة A_k في n_k نتيجة وكانت هذه التجارب متنافية فإن عدد النتائج الممكنة من إحدى هذه التجارب هو $n_1 + n_2 + \dots + n_k$.

مثال (19) : لنفرض وجود ثلاثة طرق مختلفة تؤدى من المدينة A إلى المدينة B وخمسة طرق مختلفة تؤدى من المدينة A إلى المدينة C وستة طرق مختلفة تؤدى من المدينة A إلى المدينة D فأوجد عدد الطرق التي تؤدى من المدينة A إلى إحدى المدن الثلاث ؟

الحل :

عدد الطرق التي تؤدى من المدينة A إلى إحدى المدن الثلاث هو

$$n_1 + n_2 + n_3 = 3 + 5 + 6 = 14$$

غالباً يتم السحب بدون إعادة (إحلال) بمعنى أن العنصر الذي يتم سحبه (اختياره) لا تتم إعادته قبل السحبة القادمة . وقد يكون السحب مع مراعاة الترتيب أو دون مراعاة الترتيب وذلك باستخدام التباديل والتوافيق التي تلعب دوراً هاماً في مجال حصر جميع النتائج الممكنة للتجربة العشوائية ، وسوف نتعرض هنا إلى تعريف هذين المفهومين .

III - التباديل Permutations

تعريف (17) : التباديل هي عدد الطرق المختلفة التي يمكننا بها اختيار k عنصر من n من العناصر حيث $n \geq k$ ، مع مراعاة الترتيب في كل حالة اختيار ويرمز للتباديل بالرمز $P_{n,k}$ وبالتالي فإن عدد تباديل n من العناصر المميزة المأخوذة k في كل مرة هو :

$$P_{n,k} = \frac{n!}{(n-k)!}$$

حيث الرمز ! يدل على مضروب العدد .

مثال (20) : بكم طريقة يمكن اختيار مدير لإدارة ومساعده إذا علمت أن عدد الموظفين بتلك الإدارة 12 موظفاً ؟

الحل :

حيث إنه يراعى الترتيب هنا والسحب بدون إعادة ، وعليه باستخدام التباديل نجد أن عدد الطرائق التي يمكن بها اختيارهما هو :

$$P_{12,2} = \frac{12!}{10!} = (12)(11) = 132$$

نتيجـة (6) : عدد تباديل n من العناصر المميزة المأخوذة سوية هو :

$$P_{n,n} = n(n-1)(n-1)(n-2) \ \ldots \ (2)(1) = n!$$

الحظ أن $n!$ يقرأ مضروب n وإن $1 = 0!$ و $1! = 1$

مثال (21) : بكم طريقة يمكن تنظيم ستة كتب مختلفة في رف مكتبة ما ؟

الحل :

من النتيجة (6) فإن عدد الطرق التي يمكن بها تنظيم الكتب هو : $P_{6,6} = 6! = 720$

مثال (22) : لنفرض أنه لدينا مجموعة بها k شخصاً ونود تحديد احتمال أن ـ على الأقل ـ اثنين منهم مولودان في نفس اليوم والشهر ، ولكن ليس بالضرورة في نفس السنة .

الحل :

لحل مثل هذه المسألة يجب أن نفترض أنه لا توجد علاقة بين تواريخ الميلاد لهذه المجموعة (أي لا توجد توائم) وأن لكل يوم من أيام السنة (على افتراض أن السنة 365 يوماً) نفس الفرصة بأن يكون تاريخ ميلاد أي شخص في هذه المجموعة . وحيث إنه سيكون هناك 365 تاريخ ميلاد ممكن لكل شخص في هذه المجموعة وعليه فإن فراغ العينة يحتوي على 360^k نتيجة ممكنة جميعها لها نفس الفرصة في الظهور . علاوة على ذلك فإن عدد نتائج فراغ العينة التي تمثل أن جميع تواريخ الميلاد (k) ستكون مختلفة هو $P_{365,k}$ وذلك لأن الشخص الأول من الممكن أن يكون له تاريخ ميلاد في أي يوم من 365 يوماً والشخص الثاني من الممكن أن يكون له تاريخ ميلاد في أي يوم من 364 يوماً وهكذا .إذن احتمال أن يكون لجميع الأشخاص

(k) تاريخ ميلاد مختلف هو $\frac{P_{365,k}}{365^k}$ وعليه فإن احتمال P ، أن يكون لاثنين منهما على الأقل نفس تاريخ الميلاد هو :

$$P = 1 - \frac{P_{365,k}}{365^k} = 1 - \frac{(365)!}{(365-k)!\,365^k}$$

والجدول الآتي يتضمن قيما عددية للاحتمال P لقيم مختلفة من k :

k	5	15	25	50
P	0.027	0.253	0.569	0.970

ونلاحظ من خلال هذا الجدول أنه كلما زاد عدد الأشخاص بالمجموعة كلما زاد احتمال أن يكون لاثنين منهما على الأقل نفس تاريخ الميلاد .

IV - التوافيق : Combinations

تعريف (18) : التوافيق هي عدد الطرق المختلفة التي يمكننا بها اختيار k عنصر من n من العناصر حيث أن $k \le n$ مع عدم مراعاة الترتيب في كل حالة اختيار . ويرمز للتوافيق بالرمز C_k^n أو $\binom{n}{k}$ حيث

$$\binom{n}{k} = \frac{n!}{n!(n-k)!}$$

ويسمى $\binom{n}{k}$ بمعامل ذي الحدين (binomial coefficient) وذلك لأنه يمثل معامل x^k في مفكوك ذي الحدين التالي $(1+x)^n = \sum_{k=0}^{n} \binom{n}{k} x^k$ وفي بعض الأحيان يطلق عليه عدد توافيق k من العناصر المختارة من n عنصر .

فمثلاً إذا كانت المجموعة تتضمن أربعة عناصر مميزة a , b , c , d ونود اختيار مجموعة جزئية تتضمن عنصرين من العناصر بدون مراعاة للترتيب في كل اختيار يكون عدد الاختيارات الممكنة هو : $\binom{4}{2} = \frac{4!}{2!(4-2)!} = 6$ وهي

{ a , b } , { a , c } , { a , d } , { b , c } , { b , d } , { c , d }

نلاحظ أنه عند مراعاة الترتيب تكون المجموعتان الجزئيتان {a , b} , {b , a} مختلفتين ولا فرق بينهما في حالة عدم مراعاة الترتيب وعليه نأخذ احدهما فقط.

نتيجة (7): إذا كانت n ، k فإن $\binom{n}{k}=0$ ، وحيث أن $\binom{n}{0}=\binom{n}{n}=1$ فعليه يتضح جلياً أن:

$$\binom{n}{k}=\binom{n}{n-k} \qquad , k = 0,1,2,\ldots,n$$

مثال (23): يتكون فصل دراسي من 10 طلاب و 8 طالبات. فإذا تم اختيار 5 منهم بطريقة عشوائية ما احتمال P، أن يكون من بينهم طالبتان؟

الحل:

نفرض أن الحدث A يرمز إلى أن اللجنة تتضمن طالبتان.

وحيث أنه يوجد 18 طالبا وطالبة في الفصل فإن عدد طرائق اختيار 5 من بينهم هو $\binom{18}{5}$ لكل منها نفس الاحتمال وأيضا عدد الطرائق التي يمكن بها اختيار طالبان من بين 10 طلاب هو $\binom{10}{2}$، وعدد الطرائق التي يمكن بها اختيار ثلاثة طالبات هو $\binom{8}{3}$. وحيث أن مع كل طريقة من الطرائق $\binom{10}{2}$ يمكن أن يحدث $\binom{8}{5}$ من الطرائق وعليه فإن عدد طرائق حدوثهما معاً هو $\binom{10}{2}\binom{8}{3}$. وبالتالي فإن الاحتمال المطلوب هو:

$$P(A)=\frac{\binom{10}{2}\binom{8}{3}}{\binom{18}{5}}=\frac{10!13!}{2!3!18!}=0.29$$

مثال (24): إذا ألقيت قطعة معدنية نقود متزنة 8 مرات فما احتمال الحصول على ثلاث صور بالضبط؟ وما احتمال الحصول على ثلاث صور أو اقل؟

الحل:

بفرض أن A يرمز لحدث الحصول على ثلاثة صور ، و B يرمز لحدث الحصول على ثلاثة صور أو اقل .

حيث إن عدد النتائج الممكنة بفراغ العينة هو 2^8 ، ولكل نتيجة من هذه النتائج نفس الفرصة في الظهور ، ولذلك فإن عدد النتائج التي تحتوى على ثلاث صور سوف يكون مساوياً لعدد الترتيبات المختلفة التي يمكن تكوينها بثلاث صور وخمس كتابات ، وحيث إن هذا العدد هو $\binom{8}{3}$ وعليه فإن :

$$P(A)=\frac{\binom{8}{3}}{2^8}=\frac{56}{256}=0.22$$

وبصفة عامة عدد الحالات التي يمكن فيها الحصول على k صورة هو $\binom{8}{k}$ حيث

k = 0 ، 1 ، 2 ، . . . ، 8 وبالتالي فإنه بتطبيق قاعدة الجمع نجد أن :

$$P(B)=\frac{\binom{8}{0}+\binom{8}{1}+\binom{8}{2}+\binom{8}{3}}{2^8}=\frac{1+8+28+56}{256}=0.36$$

النظرية الآتية تساعد في تحديد عدد الطرائق التي يمكن بها تقسيم n من العناصر المميزة إلى k من المجوعات المختلفة ($k \geq 2$) بحيث أن كل مجموعة تحتوي على n_i من العناصر المتشابهة أو المكررة ، حيث

$$\sum_{i=1}^{k} n_i = n$$

نظرية (6) : إذا كانت n_1 ، n_2 ، ... ، n_k أعداداً صحيحة وغير سالبة بحيث أن $\sum_{i=1}^{k} n_i = n$ فإن عدد الطرائق التي يمكن بها تجزئة المجموعة المكونة من n عنصراً إلى k مجموعة جزئية بحيث تحتوي المجموعة i على n_i من العناصر المتشابهة لجميع قيم i حيث $i=1,2,\cdots,k$ هو :

$$\binom{n}{n_1, n_2, \ldots, n_k} = \frac{n!}{n_1! \, n_2! \, \ldots \, n_k}$$

مثال (25) : ما هي عدد الطرائق التي يمكن بها تقسيم 15 طالباً بحيث يكون هناك ثلاثة طلاب تقديرهم A وستة تقديرهم C وأربعة تقديرهم D .

الحل :

من النظرية (8) يكون عدد الطرائق التي يمكن بها توزيع التقديرات مساو إلى :

$$\binom{15}{3, 2, 6, 4} = \frac{15!}{3!2!6!4!} = 6306300$$

مثال (26) : إذا تم إلقاء 12 مكعب نرد متزنة معاً مرة واحدة ، ما احتمال ظهور كل رقم من الأرقام الستة مرتين ؟

الحل :

حيث إن $n_1 = n_2 = n_3 = n_4 = n_5 = n_6 = 2$ ، وعدد النقاط بفراغ العينة هو 6^{12}

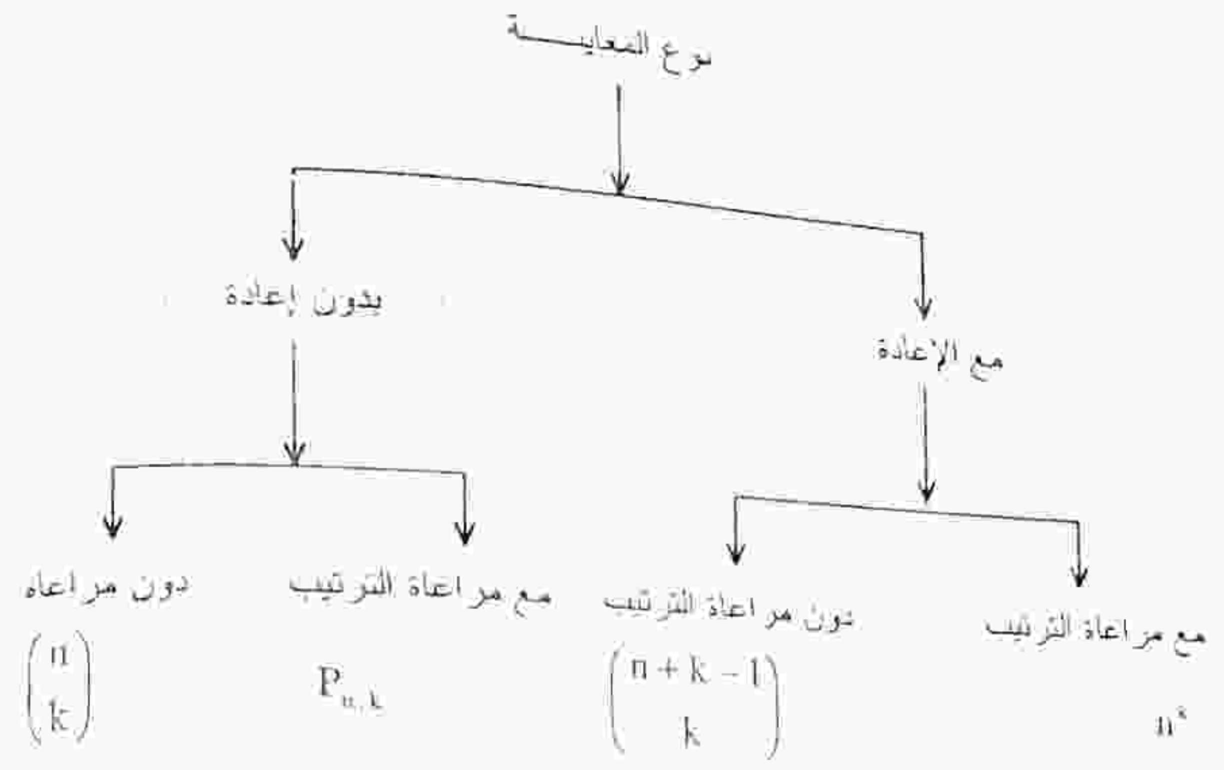
و إن عدد الطرائق التي يمكن أن يظهر فيها أي رقم مرتين من النظرية (6) هو $\frac{12!}{(2!)^6}$

وعليه فإن الاحتمال المطلوب ، P ، هو :

$$P = \frac{12!/(2!)^6}{6^{12}} = \frac{12!}{(2!)^6 \, 6^{12}} \cong 0.0003$$

وأخيراً يمكن تلخيص طرائق عد عناصر فراغ العينة السابق شرحها كما يلي :

يعتمد تحديد عدد الطرائق التي يمكن بها اختيار k من العناصر من بين n عنصراً مميزاً على الإعادة من ناحية والترتيب من ناحية أخرى وذلك كما هو موضح في الشكل التالي :



مثال (27): إذا طبعت موظفة بأحد المكاتب n رسالة وطبعت العناوين المناظرة لها على n ظرفاً ثم وضعت هذه الرسائل في الظروف بطريقة عشوائية ، فما احتمال أن رسالة واحدة على الأقل وضعتها في الظرف الصحيح ؟

الحل :

إذا فرضنا أن الحدث A_i ، $i = 1, 2, \ldots, n$ يمثل أن الرسالة i وضعت في الظرف الصحيح وعليه فإن الاحتمال المطلوب هو $P_n = P(\bigcup_{i=1}^{n} A_i)$ والذي سيتم حسابه باستخدام نتيجة (1) . حيث إن الرسائل وضعت في الظروف بطريقة عشوائية وعليه فإن احتمال أن أي رسالة معينة سوف توضع في الظرف الصحيح هو $\frac{1}{n}$ ، أي أن :

$$P(A_i) = \frac{1}{n} \qquad , \; i = 1, 2, \ldots, n$$

$$\Rightarrow \quad \sum_{i=1}^{n} P(A_i) = n \cdot \frac{1}{n} = 1$$

بالإضافة إلى ذلك ، حيث إنه من الممكن أن الرسالة الأولى وضعت في الظرف الصحيح والرسالة الثانية من الممكن أن توضع في أي واحد من الظروف (n -1) الأخرى ، وبالتالي فإن احتمال أن الرسالة الأولى والرسالة الثانية تم وضعهما في الظرفين الصحيحين هو :

$$P(A_1 \cap A_2) = \frac{1}{n(n-1)}$$

وبالمثل احتمال أن أي رسالتين i و j (j ≠ i) تم وضعهما في الظرفين الصحيحين هو :

$$P(A_i \cap A_j) = \frac{1}{n(n-1)}$$

وعليه فإن :

$$\sum_{i<j} P(A_i \cap A_j) = \binom{n}{2} \frac{1}{n(n-1)} = \frac{1}{2!}$$

وعلى غرار ذلك يكون احتمال أن أي ثلاث رسائل i، j ، k (i < j < k) يتم وضعهم في الظروف الصحيحة هو :

$$P(A_i \cap A_j \cap A_k) = \frac{1}{n(n-1)(n-2)}$$

وعليه فإن :

$$\sum_{i<j<k} P(A_i \cap A_j \cap A_k) = \binom{n}{3} \frac{1}{n(n-1)(n-2)} = \frac{1}{3!}$$

وبنفس الأسلوب يمكن أن نستنتج أن احتمال وضع جميع الرسائل في الظروف الصحيحة هو :

$$P(A_1 \cap A_2 \cap A_3 \cap \cdots \cap A_n) = \frac{1}{n(n-1)(n-2)\ \ldots\ 2.1} = \frac{1}{n!}$$

وعليه يكون احتمال وضع رسالة واحدة على الأقل في الظرف الصحيح هو :

$$P_n = P(\bigcup_{i=1}^{n} A_i) = 1 - \frac{1}{2!} + \frac{1}{3!} - \frac{1}{4!} + \ \ldots\ + (-1)^{n-1}\frac{1}{n!}$$

ومن نظرية بالتفاضل وعندما $n \to \infty$ فإن الطرف الأيمن بالمعادلة أعلاه ستكون نهايته كالآتي:

$$\lim_{n \to \infty} P_n = 1 - \frac{1}{e} = 0.63212$$

حيث $e = 2.71828$. ألحظ أنه إذا كانت $n \geq 7$ فإن قيم P_n تتناقص وتكون لها القيمة 0.6321 وعليه فإن الاحتمال يبقى ثابتاً لقيم $n \geq 7$.

2 - 7 الاحتمال الشرطي Conditional Probability

2.7 . 1 -مقدمـــة

في هذا البند سوف ندرس الطريقة التي يتغير بها احتمال وقوع الحدث A بعد توفر معلومات عن وقوع حدث آخر وليكن B مثلاً . هذا الاحتمال الجديد للحدث A يسمى الاحتمال الشرطي للحدث A إذا علم وقوع الحدث B . ويرمز له بالرمز $P(A|B)$ ويقرأ :

* احتمال وقوع الحدث A بمعلومية الحدث B *.

شكل (12) : نتائج في الحدث B وتنتمي ايضاً للحدث A

إذن إذا علم وقوع الحدث B فإننا نعلم بأن نتيجة التجربة محتواة في B . وعليه لتقويم احتمال إمكانية حدوث الحدث A يجب دراسة مجموعة النتائج الموجودة في B والتي أيضاً ستظهر في الحدث A كما في شكل (12) ، وفي الحقيقة أن هذه المجموعة متمثلة في التقاطع $A \cap B$.إذن من الطبيعي تعريف الاحتمال الشرطي $P(A|B)$ بأنه نسبة $P(A \cap B)$ إلى الاحتمال الكلي $P(B)$. فاحتمال حدوث الحدث A بشرط وقوع الحدث B يسمى أحياناً بالاحتمال النسبي للحدث A ، لأنه عبارة عن احتمال حدوث الحدث A بالنسبة لعدد الحالات التي حدث فيها الحدث B . وعليه يمكن صياغة تعريف الاحتمال الشرطي كما يلي :

تعريف (19) : الاحتمال الشرطي Conditional Probability

إذا كان A و B حدثين فإن الاحتمال الشرطي للحدث A إذا علم حدوث الحدث B يعرف كما يلي:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad , \; P(B) > 0$$

وبالمثل يكون الاحتمال الشرطي للحدث B إذا علم حدوث الحدث A هو :

$$P(B|A) = \frac{P(A \cap B)}{P(A)} \quad , \; P(A) > 0$$

أن للاحتمال الشرطي $P(A|B)$ معنى بسيط يمكن تفسيره من خلال مفهوم التكرار النسبي للاحتمال . فعلى ضوء هذا المفهوم إذا أعيدت التجربة عدداً كبيراً من المرات فإن نسبة التكرار التي سيظهر فيها الحدث B تساوي تقريباً $P(B)$ ، وأن نسبة التكرار التي سيظهر فيها

الحدثين A و B تساوي تقريباً $P(A \cap B)$. إذن من بين جميع التكرارت التي ظهر فيهـا الحـدث B فإن نسبة التكرارات التي يظهر فيها الحدث A تقريباً مساوية إلى :

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

ولمزيد من التوضيح نفترض أن :

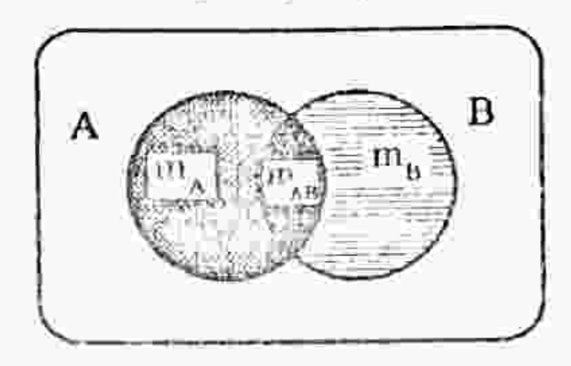
العدد الكلي للحالات الممكنة (عدد عناصر فراغ العينة) $= n$

عدد الحالات التي يتحقق فيها الحدث A فقط $(A - B)$ $= m_A$

عدد الحالات التي يتحقق فيها الحدث B فقط $(B - A)$ $= m_B$

عدد الحالات التي يتحقق فيها الحدثان A و B معاً $(A \cap B)$ $= m_{AB}$

ويمكن توضيح ذلك بيانياً كما في شكل (13) التالي :



شكل (13) : عدد الحالات التي تحقق الحدث A شريطة وقوع الحدث B .

إذا علمنا أن الحدث B قد حدث ، فإن احتمال حدوث الحدث A بمعلومية الحدث B هو :

$$P(A|B) = \frac{m_{AB}}{m_B + m_{AB}}$$

وبقسمة كل من البسط والمقام على n نحصل على الآتي :

$$P(A|B) = \frac{\frac{m_{AB}}{n}}{\frac{m_B + m_{AB}}{n}}$$

ولكن

$$P(A \cap B) = \frac{m_{AB}}{n} \quad , \quad P(B) = \frac{m_B + m_{AB}}{n}$$

ومنها

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad , \quad P(B) > 0$$

والسؤال المهم هو هل يمكن تبرير قولنا إن $P(A|B)$ يمثل احتمالاً ؟ وبعبارة أخرى هل أن $P(A|B)$ تفي بمسلمات الاحتمال ؟ الجواب هو نعم . إن $P(A|B)$ تفي بمسلمات الاحتمالات السالفة الذكر في بند (2 - 5) وذلك للأسباب التالية :

(I) من تعريف الاحتمال الشرطي نجد أن :

$$P(A|B)=\frac{P(A\cap B)}{P(B)} \quad , \; P(B)>0$$

ولما كان

$$0 \le P(A\cap B) \le P(B)$$

وعليه فإن :

$$0 \le P(A|B) \le 1$$

(II)

$$P(\Omega|B)=\frac{P(\Omega\cap B)}{P(B)}=\frac{P(B)}{P(B)}=1$$

(III) إذا كانت $A_1 , A_2 , \ldots , A_n$ تمثل متوالية من الأحداث المتنافية فإن :

$$P\left(\bigcup_{i=1}^{\infty} A_i|B\right)=\frac{P\left(\bigcup_{i=1}^{\infty} A_i\cap B\right)}{P(B)}=\frac{P\left(\bigcup_{i=1}^{\infty} A_i\cap(A_i\cap B)\right)}{P(B)}$$

$$=\frac{\sum_{i=1}^{\infty} P(A_i\cap B)}{P(B)}=\sum_{i=1}^{\infty} P(A_i|B)$$

مثال (28) : إذا تم إلقاء مكعب نرد متزن (منتظم) مرتين ولوحظ أن مجموع الرقمين (T) عدد زوجي ، ما احتمال أن يكون T أقل من 8 ؟

الحل :

نفرض أن الحدث A هو أن يكون $T<8$ والحدث B هو أن يكون T زوجياً ، انظر الجدول أعلاه ، فإننا نحصل على :

$$P(A\cap B)=\frac{2+4+6}{36}=\frac{12}{36}=\frac{1}{3}$$

$$P(B)=\frac{2+4+6+4+2}{36}=\frac{18}{36}=\frac{1}{2}$$

نتيجة المكعب I / نتيجة المكعب II	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

وعليه فإن :
$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{2}{3}$$

مثال (29) : بفرض أن الجدول التالي يبين عدد طلاب الفصل الدراسي الأول في كلية العلوم بجامعة الفاتح في العام الجامعي 98 - 1999 ف مصنفين حسب التخصص (المجموعات) والجنس .

المجموعه / الجنس	علوم رياضية	علوم الحياة	علوم طبيعية	المجموع
ذكور	200	100	100	400
إناث	400	300	100	800
المجموع	600	400	200	1200

فإذا تم اختيار واحد منهم بطريقة عشوائية أوجد احتمال حدوث الأحداث التالية :

أ - أن يكون الشخص من مجموعة العلوم الرياضية وذكر .

ب - أن يكون الشخص من مجموعة علوم الحياة علماً بأنها طالبة .

جـ - أن يكون ذكر بشرط أن يكون في مجموعة العلوم الرياضية .

الحل :

نفترض أن :

حدث اختيار شخص من مجموعة العلوم الرياضية = A

حدث اختيار شخص من مجموعة علوم الحياة = B

حدث اختيار شخص من مجموعة العلوم الطبيعية = C

حدث اختيار ذكر = M

حدث اختيار أنثى = F

أ - $P(M\cap A)=\frac{200}{1200}=\frac{1}{6}$

ب - $P(B|F)=\frac{300}{800}=\frac{3}{8}$

جـ - $P(M|C)=\frac{100}{200}=\frac{1}{2}$

ملحوظـة : إذا توفرت معلومات عن وقوع الحدث A فهذا لا يعني بالضرورة أن فرصة حدوث الحدث B ستكون كبيرة ، وفي الواقع إذا كان $A\cap B=\varnothing$ فإن :

$$P(B\,|\,A)=0\le P(B)$$

وهذا يعني أن معلوماتنا عن الحدث A توحي لنا بأن الحدث B لا يمكن حدوثه . ومن ناحية أخرى ، إذا كانت $A\subseteq B$ فإن حدوث الحدث B لا يؤدي إلى التقليل من فرصة حدوث الحدث A . وفي الواقع بما أن $A\cap B=B$ فإن :

$$P(B\,|\,A)=\frac{P(B)}{P(A)}\ge P(B)$$

وإذا كانت $A\supseteq B$ فإنه بديهي $P(B\,|\,A)=1$.

وأخيراً سوف نعرض مجموعة النتائج التالية وذلك على افتراض أن B حدثاً بحيث $P(B)>0$:

نتيجـــة (8) : $P(\phi|B)=0$

البرهان :

$$P(\phi|B)=\frac{P((\phi\cap B)}{P(B)}=\frac{P(\phi)}{P(B)}=0$$

نتيجـــة (9) : إذا كان A و B حدثين فإن : $P(A'|B)=1-P(A|B)$

البرهان :

$$P(A'|B)=\frac{P(A'\cap B)}{P(B)}=\frac{P(B-A)}{P(B)}$$

$$=\frac{P(B)-P(A\cap B)}{P(B)}$$

$$=\frac{P(B)}{P(B)}-\frac{P(A\cap B)}{P(B)}$$

$$=1-P(A|B)$$

نتيجـــة (10) : إذا كان A_1 , A_2 حدثين بحيث $A_1\subset A_2$ فإن :

$$P(A_1|B)\le P(A_2|B)$$

البرهان :

حيث إن $A_1\subset A_2 \quad \Leftarrow \quad P(A_1)\le P(A_2) \quad \Leftarrow$

$$P(A_1|B)=\frac{P(A_1\cap B)}{P(B)}\le\frac{P(A_2\cap B)}{P(B)}=P(A_2|B)$$

مثال (30) : في اختبار نهاية الفصل الدراسي وجد أن 40 %من الطلبة نجحوا في مادة الإحصاء و 25 % نجحوا في مادة الحاسوب و 20 % نجحوا في مادتي الإحصاء والحاسوب. فإذا تم اختيار أحد الطلاب عشوائياً وكان الحدث A يمثل نجاح الطالب في الإحصاء والحدث B يمثل نجاحه في الحاسوب فأوجد احتمال حدوث الأحداث التالية :

أ – نجاح الطالب في الحاسوب إذا علمنا أنه نجح في الإحصاء .

ب – نجاح الطالب في الإحصاء إذا علمنا أنه نجح في الحاسوب .

جـ – نجاحه في الإحصاء إذا علمنا رسوبه في الحاسوب .

د – رسوب الطالب في الحاسوب بشرط رسوبه في الإحصاء .

الحل :

من المعطيات نجد أن :

$$P(A)=0.40 \quad , \quad P(B)=0.25 \quad , \quad P(A\cap B)=0.20$$

وبالتالي فإن :

أ - احتمال نجاح الطالب في الحاسوب إذا علمنا أنه نجح في الإحصاء هو :

$$P(B|A)=\frac{P(A\cap B)}{P(A)}=\frac{0.20}{0.40}=0.50$$

ب - احتمال نجاح الطالب في الإحصاء إذا علمنا أنه نجح في الحاسوب هو :

$$P(A|B)=\frac{P(A\cap B)}{P(B)}=\frac{0.20}{0.25}=0.80$$

جـ - احتمال نجاحه في الإحصاء إذا علمنا رسوبه في الحاسوب هو :

$$P(A|B')=\frac{P(A\cap B')}{P(B')}=\frac{P(A)-P(A\cap B)}{1-P(B)}=\frac{0.40-0.20}{1-0.25}=\frac{0.20}{0.75}\cong 0.27$$

د - احتمال رسوب الطالب في الحاسوب بشرط رسوبه في الإحصاء هو :

$$P(B'|A')=\frac{P(A'\cap B')}{P(A')}=\frac{P\left((A\cup B)'\right)}{P(A')}=\frac{1-P(A\cup B)}{1-P(A)}$$

$$=\frac{1-\left[P(A)+P(B)-P(A\cap B)\right]}{1-P(A)}$$

$$=\frac{1-(0.40+0.25-0.20)}{1-0.40}\approx 0.92$$

2 . 7 . 2 قاعدة الضرب للاحتمالات الشرطية :

The Multiplication Rule For Conditional Probabilities

في بعض الأحيان نجد أنه من المناسب حساب $P(A\cap B)$ وذلك بتطبيق الصيغة التالية التي تم استنباطها من تعريف الاحتمال الشرطي بضرب الطرفين في الوسطين :

$$P(A\cap B)=\begin{cases}P(A).P(B|A) & ,P(A)>0\\ P(B).P(A|B) & ,P(B)>0\end{cases}$$

مثال (31) : لنفرض أننا سنختار كرتين بطريقة عشوائية وبدون إعادة ، من صندوق به r كرة حمراء و b كرة زرقاء . احسب احتمال أن الكرة المختارة الأولى حمراء و الثانية زرقاء .

الحل :

نفترض أن الحدث A يمثل أن الكرة الأولى المختارة حمراء ، و الحدث B يمثل أن الكرة الثانية المختارة زرقاء ، وعليه فإن :

$$P(A) = \frac{r}{r+b}$$

علاوة على ذلك إذا كان الحدث A قد تحقق فإن احتمال أن تكون الكرة المسحوبة ثانياً زرقاء هو

$$P(B \mid A) = \frac{b}{r+b-1}$$

وعليه فإن الاحتمال المطلوب هو :

$$P(A \cap B) = P(A) \,.\, P(B \mid A) = \frac{r\,b}{(r+B)(r+b-1)}$$

مثال (32) : يحتوي صندوق على عشرة مصابيح كهربائية من بينها أربعة معيبة ، فإذا سحب مصباحان عشوائياً الواحد تلو الآخر وبدون إعادة فاحسب احتمال حدوث الأحداث التالية :

أ - أن يكون المصباحان معيبين .　　　　ب - أن يكون المصباحان صالحين .

جـ - أن يكون الأول صالحاً والثاني معيباً .　　　　د - أحدهما على الأقل صالح .

الحل :

أ - إذا كان الحدث A يمثل أن المصباح الأول معيب والحدث B يمثل أن المصباح الثاني معيب فإن الاحتمال المطلوب هو :

$$P(A \cap B) = P(A)\,P(B|A) = \frac{4}{10} \cdot \frac{3}{9} = \frac{12}{90} \cong 0.13$$

ب - إذا كان الحدث C يمثل أن المصباح الأول صالح والحدث D يمثل أن المصباح الثاني صالح فإن احتمال أن يكون المصباحان صالحين هو :

$$P(C \cap D) = P(C) \,.\, P(D \mid C) = \frac{6}{10} \cdot \frac{5}{9} = \frac{30}{90} \cong 0.33$$

جـ - إذا رمزنا لحدث أن المصباح الأول صالح بالرمز E ولحدث أن المصباح الثاني معيب بالرمز F فإن الاحتمال المطلوب هو :

$$P(E \cap F) = P(E) \,.\, P(F \mid E) = \frac{6}{10} \cdot \frac{4}{9} = \frac{24}{90} \cong 0.27$$

د - واحتمال أن أحدهما على الأقل صالح هو :

$$P\left[(C \cap D)' \right] = 1 - P(C \cap D) = 1 - \frac{12}{90} = \frac{78}{90} \cong 0.87$$

الحظ أنه إذا كان السحب بدون إعادة مع مراعاة الترتيب فإن احتمال أن يكون المصباحان معيبين هو :

$$P(A \cap B) = \frac{P_{4,2}}{P_{10,2}} = \frac{4 \times 3}{10 \times 9} = \frac{12}{90} \cong 0.13$$

بينما احتمال أن يكون المصباحان صالحين هو :

$$P(C \cap D) = \frac{P_{6,2}}{P_{10,2}} = \frac{5 \times 6}{10 \times 9} = \frac{30}{90} \cong 0.33$$

وأن احتمال أن يكون الأول صالحاً والثاني معيباً هو :

$$P(C \cap B) = \frac{P_{6,1} \times P_{4,1}}{P_{10,2}} = \frac{6 \times 4}{90} = \frac{24}{90} \cong 0.27$$

إن الفكرة التي تم تطبيقها في هذين المثالين يمكن تعميمها في النظرية أدناه إلى أي عدد محدود من الأحداث .

نظريــة (7) : إذا كانت $A_1, A_2, \cdots, A_n$ تمثل احداثاً بحيث أن $P(\bigcap_{i=1}^{k} A_i) > 0$ و $k = 1, 2, \ldots, n-1$ فإن :

$$P\left(\bigcap_{i=1}^{n} A_i\right) = P(A_1) P(A_2 | A_1) P(A_3 | A_1 \cap A_2) \cdots P(A_n | A_1 \cap A_2 \cap \cdots \cap A_{n-1})$$

البرهان :

إن الطرف الأيمن في المعادلة أعلاه يساوي :

$$P(A_1) \cdot \frac{P(A_1 \cap A_2)}{P(A_1)} \cdot \frac{P(A_1 \cap A_1 \cap A_2)}{P(A_1 \cap A_2)} \cdots \frac{P(A_1 \cap A_2 \cap \cdots \cap A_{n1})}{P(A_1 \cap A_2 \cap \cdots \cap A_{n-1})}$$

حيث جميع هذه الحدود تختصر مع بعضها البعض عدا الحد الأخير بالبسط وهو : $P(A_1 \cap A_2 \cap \cdots \cap A_n)$ والذي يمثل الطرف الأيسر في تلك المعادلة .

الحظ أنه يمكن البرهان بطريقة أخرى وذلك من خلال ملاحظة أن :

$$A_1 \supseteq (A_1 \cap A_2) \supseteq \cdots \supseteq (A_1 \cap A_2 \cap \cdots \cap A_{n-1})$$

لذلك فإن
$$P(A_1) \geq P(A_1 \cap A_2) \geq \cdots \geq P(\bigcap_{i=1}^{n-1} A_i) > 0$$

وعليه فإن الاحتمالات الشرطية في هذه النظرية معرفة وبذلك يمكن البرهان باستخدام الاستنتاج الرياضي وتعريف الاحتمال الشرطي (قانون ضرب الاحتمالات) .

مثال (33) : سحبت أربع كرات وبدون إعادة من صندوق به r كرة حمراء و b كرة زرقاء ، ما احتمال الحصول على متوالية من النتائج الآتية :

الكرة الأولى حمراء والثانية زرقاء والثالثة حمراء والرابعة زرقاء ؟

الحل :

لنفرض أن R_i تمثل حدث الحصول على كرة حمراء في السحبة i = 1 ، 3 وإن الحدث B_j يمثل حدث الحصول على كرة زرقاء في السحبة j = 2 ، 4 . إذن الاحتمال المطلوب هو :

$$P(R_1 \cap B_2 \cap R_3 \cap B_4) = P(R_1).P(B_2|R_1).P(R_3|R_1 \cap B_2).P(B_4|R_1 \cap B_2 \cap R$$

$$= \frac{r}{r+b} \cdot \frac{b}{r+b-1} \cdot \frac{r-1}{r+b-2} \cdot \frac{b-1}{r+b-3}$$

مثال (34) : يحتوى صندوق على 8 كرات حمراء و6 كرات بيضاء فإذا تم اختيار ثلاث كرات بطريقة عشوائية ، ما احتمال أن تكون الكرات جميعها حمراء ؟

الحل :

يمكن حل هذا المثال بطرقتين ، الأولى هي أن الكرات تم سحبها الواحدة وراء الأخرى وافترضنا أن الحدث A يمثل أن الكرة الأولى حمراء والحدث B يمثل أن الكرة الثانية حمراء والحدث C يمثل أن الكرة الثالثة حمراء . فإن الاحتمال المطلوب هو :

$$P(A \cap B \cap C) = P(A).P(B|A).P(C|A \cap B)$$

$$= \frac{8}{14} \cdot \frac{7}{13} \cdot \frac{6}{12} = \frac{2}{13} \cong 0.15$$

أما الطريقة الثانية فهي أن الكرات الثلاثة تم سحبها مع بعض مرة واحدة وفي مثل هذه الحالة يكون الاحتمال المطلوب ، p ، هو :

$$p = \frac{\binom{8}{3}}{\binom{14}{3}} = \frac{8 \times 7 \times 6}{14 \times 13 \times 12} = \frac{2}{13} \cong 0.15$$

Partition of Sample Space تجزئة فراغ العينة

تعريف (20) : تجزئة فراغ العينة لجميع قيم $\{A_i\}$ الأحداث فإنه يقال بأن متتالية من $A_1, A_2, \ldots, A_n$ تمثل ... لفراغ العينة إذا وفقط إذا (iff) تحقق الشرطان التاليان : $i = 1, 2, \ldots$ هي تجزئة لفراغ العينة

$$(1) \quad \bigcup_{i=1}^{\infty} A_i = \Omega$$

$$(2) \quad A_i \cap A_j = \phi \quad \forall\, i, j = 1, 2, \ldots,\; i \neq j$$

مثال (35) : تجربة عشوائية تتألف من رمي قطعة نقود معدنية مرة واحدة . عندئذ يكون :

$$\Omega = \{H, T\}$$

ومن ذلك نحصل على أن المجموعة $\{\{H\}, \{T\}\}$ هي تجزئة لفراغ العينة لأن :

$$(I) \quad \{H\} \cup \{T\} = \Omega$$

$$(II) \quad \{H\} \cap \{T\} = \phi$$

مثال (36) : تجربة عشوائية تتألف من رمي مكعب نرد مرة واحدة . عندئذ يكون :

$$\Omega = \{1, 2, 3, 4, 5, 6\}$$

فإن المجموعة $\{\{1\}, \{2, 3\}, \{4, 5, 6\}\}$ تمثل تجزئة لفراغ العينة وذلك لأن :

$$(1) \qquad \{1\} \cup \{2, 3\} \cup \{4, 5, 6\}$$

$$(2) \qquad \{1\} \cap \{2, 3\} \cap \{4, 5, 6\}$$

كما أن المجموعتين التاليتين يمثل كل منهما تجزئة لفراغ عينة التجربة أعلاه :

$$\{\{1, 2\}, \{3, 4\}, \{5, 6\}\}$$

$$\{\{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{6\}\}$$

نظريـــة (8) نظرية الاحتمال الكلي Total Probability Theorem

[illegible] $A_1, A_2, \ldots$ تجزئة لفراغ العينة Ω [illegible] $P(A_i) > 0$ [illegible]
[illegible] B [illegible] فراغ العينة [illegible]

$$P(B) = \sum_{i=1}^{\infty} P(B|A_i)\, P(A_i)$$

البرهان

$$B = B \cap \Omega$$
$$= B \cap (A_1 \cup A_2 \cup \cdots)$$
$$= (B \cap A_1) \cup (B \cap A_2) \cup \cdots$$

الحظ أن الأحداث $B \cap A_n$ ، لجميع قيم n ، متنافية . لذلك

$$P(B) = P(B \cap A_1) + P(B \cap A_2) + \cdots$$
$$= P(B|A_1).P(A_1) + P(B|A_2).P(A_2) + \ldots$$
$$= \sum_{n=1}^{\infty} P(B|A_n).P(A_n)$$

A_1	A_2	. . .	A_i	. . . .	A_n
		B			

شكل (14) : تقاطع الحدث B مع الأحداث المجزأة لفراغ العينة

مثال (37) : إذا علمت بأنه في أحد المصانع ثلاث خطوط للإنتاج ينتج الخط الأول 50 % من إنتاج المصنع ، وينتج الخط الثاني 30 % من الإنتاج والباقي يقوم بإنتاجه الخط الثالث . ما احتمال إنتاج وحدة معيبة (defective) في المصنع ككل علماً بأن نسب الإنتاج المعيب في الخطوط الثلاثة على الترتيب هي 2 % ، 3 % ، 5 % .

الحل :

نفرض ان A_i = حدث أن الإنتاج كان من قبل الخط i ، حيث i = 1 ، 2 ، 3

D = حدث إنتاج وحدة معيبة .

وعليه فإن :

$$P(A_1) = 0.50 \quad , \quad P(A_2) = 0.30 \quad , \quad P(A_3) = 0.20$$

وأن الحدث (D / A_i) يعني الوحدة معيبة علماً بأنها منتجة من قبل الخط i ، وبذلك يكون :

$$P(D / A_1) = 0.02 \quad , \quad P(D / A_2) = 0.03 \quad , \quad P(D / A_3) = 0.05$$

وبالتالي فإن الاحتمال المطلوب هو :

$$P(D)=\sum_{i=1}^{3} P(A_i).P(D|A_i)$$
$$=P(A_1).P(D|A_1)+P(A_2).P(D|A_2)+P(A_3).P(D|A_3)$$
$$=0.50\times 0.02+0.30\times 0.03+0.20\times 0.05$$
$$=0.01+0.006+0.01=0.026$$

نتيجـــة (11) : لأي حدثين A و B يكون :
$$P(A)=P(A|B).P(B)+P(A|B').P(B')$$

مثال (38) : يحتوى الصندوق " I " على كرتين لونهما ابيض وأربع كرات حمراء بينما يحتوى الصندوق " II " على كرة بيضاء وأخرى حمراء ، فإذا سحبت كرة من الصندوق " I " ووضعت في الصندوق " II " ، ثم سحبت كرة بطريقة عشوائية من الصندوق " II " . فما احتمال أن تكون الكرة المسحوبة من الصندوق " II " بيضاء ؟

الحل :

لنفرض أن الحدث B يمثل الكرة المنقولة من الصندوق الأول إلى الصندوق الثاني بيضاء.
والحدث B' يمثل الكرة المنقولة من الصندوق الأول إلى الصندوق الثاني حمراء .
والحدث A يمثل الكرة المسحوبة من الصندوق الثاني بيضاء ، إذن من النتيجة السابقة نجد أن :
$$P(A)=P(A|B).P(B)+P(A|B').P(B')$$
$$=\frac{2}{3}\times\frac{2}{6}+\frac{1}{3}\times\frac{4}{6}=\frac{4}{9}\cong 0.44$$

نظريـــة (9) : نظرية بييز Baye ' s Theorem

إذا كانت $A_1, A_2, \cdots, A_n$ أحداث تجزئ فراغ (فضاء) العينة Ω بحيث $P(A_i)>0$ لجميع قيم i وكان B أي حدث بشرط أن $P(B)>0$ فإن :

$$P(A_i|B)=\frac{P(A_i).P(B|A_i)}{\sum_{j=1}^{n} P(A_j)P(B|A_j)} \quad , \; i=1,2,\cdots,n$$

البرهان :

لما كانت

$$P(A_i \cap B) = P(A_i)P(B|A_i)$$
$$= P(B)P(A_i|B)$$

عندئذ يكون

$$P(A_i|B) = \frac{P(A_i)P(B|A_i)}{\sum_{j=1}^{n} P(A_j)P(B|A_j)}$$

وذلك بالاستناد على نظرية الاحتمال الكلي .

إن نظرية بييز مفيدة في الإجابة على التساؤل الذي يكون من النوع التالي : إذا وقع الحدث B مثلاً فما احتمال أنه وقع بسبب الحدث A_n .

مثال (39) : في المثال رقم (38) إذا تم اختيار وحدة من الوحدات المنتجة في المصنع فوجد أنها معيبة فأي الخطوط الثلاث ترجح أنه قام بإنتاجها ؟

الحل :

من المثال السابق الحظ أن

$$P(A_1) = 0.50 \quad , \quad P(D|A_1) = 0.02$$
$$P(A_2) = 0.30 \quad , \quad P(D|A_2) = 0.03$$
$$P(A_3) = 0.20 \quad , \quad P(D|A_3) = 0.05$$

وباستخدام نظرية بييز نجد أن :

$$P(A_1|D) = \frac{P(A_1)P(D|A_1)}{\sum_{i=1}^{3} P(A_i)P(D|A_i)} = \frac{P(A_1)P(D|A_1)}{P(D)} = \frac{0.01}{0.029} \cong 0.345$$

$$P(A_2|D) = \frac{P(A_2)P(D|A_2)}{P(D)} = \frac{0.009}{0.029} \cong 0.31$$

$$P(A_3|D) = \frac{P(A_3)P(D|A_3)}{P(D)} = \frac{0.01}{0.029} \cong 0.345$$

وحيث أن احتمال أن الوحدة المعيبة كانت من إنتاج الخط الأول أو الثالث هو الأكبر ، عليه نرجح أن الخط الأول أو الخط الثالث هو الذي أنتج هذه الوحدة المعيبة .

إن الاحتمال $P(A_i)$ و $i = 1 , 2 , 3$ في هذا المثال غالباً ما يطلق عليه الاحتمال المسبق (prior probability) لحدث الوحدة المختارة التي تم إنتاجها بالآلة i ، والسبب في ذلك لأنه يمثل احتمال وقوع هذا الحدث قبل اختيار تلك الوحدة وقبل المعرفة بأنها معيبة أو سليمة . بينما يطلق على الاحتمال $P(A_i|D)$ الاحتمال اللاحق (posterior probability) للحدث الوحدة المختارة التي تم إنتاجها بالآلة i ، وذلك لأنه يمثل احتمال وقوع هذا الحدث بعد معرفة أن الوحدة المختارة معيبة .

مثال (40) : يحتوي صندوق على ثلاث بطاقات متشابهة منه بطاقة لونها أحمر من الجهتين ، وأخرى لونها أبيض من الجهتين ، وأخرى لونها أحمر من جهة وأبيض من الجهة الأخرى . وبفرض أنه تم اختيار بطاقة بطريقة عشوائية من ذلك الصندوق ، وكانت الجهة العلوية لها حمراء اللون . ما احتمال أن تكون الجهة الأخرى لونها أبيض ؟

الحل :

لنفرض أن الحدث RR يمثل أن البطاقة التي تم اختيارها لونها أحمر من الجهتين ، والحدث WW يمثل أن البطاقة التي تم اختيارها لونها أبيض من الجهتين ، والحدث RW يمثل أن البطاقة التي تم اختيارها جهة لونها أحمر وجهة لونها أبيض . وبفرض أن الحدث R يمثل أن البطاقة التي تم اختيارها كانت جهتها العلوية حمراء ، عندئذ يكون الاحتمال المطلوب هو :

$$P(RW/R) = \frac{p(RW \cap R)}{P(R)}$$

$$= \frac{P(R|RW)P(RW)}{P(R|RR)P(RR) + P(R|RW)P(RW) + P(R|WW)P(WW)}$$

$$= \frac{\frac{1}{2}(\frac{1}{3})}{1(\frac{1}{3}) + \frac{1}{2}(\frac{1}{3}) + 0(\frac{1}{3})} = \frac{1}{3}$$

2 - 8 الأحداث المستقلة Independent Events

يعتبر هذا الموضوع من أهم الموضوعات الأساسية في نظرية الاحتمال ، وهو حالة خاصة من قاعدة الضرب التي سبق الإشارة إليها . فإذا كان A و B حدثين وكان وقوع أو عدم وقوع أي منهما لا علاقة ولا تأثير له على وقوع أو عدم وقوع الحدث الأخر فإن A و B حدثان مستقلان . وبعبارة أخرى يكون من الطبيعي افتراض بأن احتمال وقوع A و B معاً يساوي حاصل ضرب احتمال وقوع كل واحد منهما على حده . إن هذه النتيجة يمكن بسهولة تبريرها بدلالة مفهوم التكرار النسبي للاحتمال ، فعلى سبيل المثال إذا كان الحدث A يمثل ظهور صورة عند إلقاء قطعة نقود معدنية متزنة مرة واحدة والحدث B يمثل ظهور العدد " 1 " أو العدد " 2 " عند إلقاء مكعب نرد متزن . وعليه فإن الحدث A سيحدث بتكرار نسبي يساوي $\frac{1}{2}$ عند إلقاء قطعة النقود عدداً كبيراً من المرات ، والحدث B سيحدث بتكرار نسبي يساوي $\frac{1}{3}$ عند إلقاء مكعب النرد عدداً كبيراً من المرات . وبالتالي فإن $P(A)=\frac{1}{2}$ و $P(B)=\frac{1}{3}$.

ولندرس الآن التجربة التي يتم فيها إلقاء مكعب النرد وقطعة النقود معاً ، فإذا أجريت هذه التجربة عدداً كبيراً من المرات فإن التكرار النسبي للحدث A (كما سبق تعريفه) سيبقى $\frac{1}{2}$ ، وذلك لأن نتائج قطعة النقود ونتائج مكعب النرد لا علاقة لهما ببعضهما البعض خلال تلك التجارب التي يحدث فيها الحدثان ، وإن التكرار النسبي للحدث B (كما سبق تعريفه) سيبقى $\frac{1}{3}$. وعليه في متوالية من هذه التجربة يكون التكرار النسبي لوقوع الحدثين A و B معاً هو $\frac{1}{2}\times\frac{1}{3}=\frac{1}{6}$ أي أن

$$P(A\cap B)=P(A)\,P(B)=\frac{1}{2}\times\frac{1}{3}=\frac{1}{6}$$

ومن خلال هذا المثال ، يمكن صياغة التعريف الرياضي التالي لاستقلالية حدثين كما يلي :

تعريف (21) : الاستقلالية The Independence

يقال بأن الحدثين A و B مستقلان إذا وفقط إذا (III) تحقق أحد الشروط الآتية :

(I) $P(A\cap B)=P(A)\,P(B)$

(II) $P(A|B)=P(A)$, $P(B)>0$

(III) $P(B|A)=P(B)$, $P(A)>0$

ولتوضيح أن الشروط الثلاثة أعلاه متكافئة فإنه يكفي التوضيح بأن : $(I) \Leftarrow (III)$ وأن $(III) \Leftarrow (II)$ وأن $(II) \Leftarrow (I)$

فإذا كان $P(A \cap B) = P(A)P(B)$ فإن :

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)P(B)}{P(B)} = P(A) \quad , P(B) > 0$$

وعليه فإن $(I) \Rightarrow (II)$.

وإذا كان $P(A|B) = P(A)$ فإن :

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)} = \frac{P(A)P(B)}{P(A)} = P(B) \quad , P(B) > 0 , P(A) > 0$$

وبالتالي فإن $(II) \Rightarrow (III)$.

وإذا كان $P(B|A) = P(B)$ فإن

$$P(A \cap B) = P(B/A)P(A) = P(B)P(A) \quad , P(A) > 0$$

وعليه فإن $(III) \Rightarrow (I)$.

من الواضح أن $P(A \cap B) = P(A)P(B)$ إذا كان $P(A) = 0$ أو $P(B) = 0$ وأن العلاقة $P(A \cap B) = P(A)P(B)$ متماثلة في A و B ، بمعنى أن

$$P(A \cap B) = P(B \cap A) = P(A)P(B) = P(B)P(A)$$

على القارئ أن يلحظ وجود اختلاف بين الأحداث المستقلة والأحداث المتنافية ، فليس من الضروري أن يتضمن أحدهما الآخر . فعلى سبيل المثال ، إذا كان A و B حدثين متنافيين وكان $P(A) > 0$ و $P(B) > 0$ فإن $P(A \cap B) = 0 \neq P(A)P(B)$ وبالتالي يكون الحدثان A و B غير مستقلين . وإذا كان A و B حدثين مستقلين ، وكان $P(A) > 0$ و $P(B) > 0$ فإن $P(A \cap B) > 0$ وبالتالي يكون الحدثان غير متنافيين . ولكن يكون الحدثان المتنافيان مستقلين إذا وفقط إذا كان $P(A)P(B) = 0$ وهذا صحيح إذا وفقط إذا كان لأحد الحدثين احتمال مساوي للصفر . وبتعبير آخر ، يكون الحدثان المتنافيان مستقلين إذا وفقط إذا كان احتمال أحدهما مساوي للصفر .

مثال (41) : بفرض أنه ألقيت قطعة نقود معدنية متزنة مرتين ، وعرفنا الحدث A بأنه حدث الحصول على وجهين متشابهين ، والحدث B حدث الحصول على صورة واحدة على الأقل ، والحدث C حدث الحصول على صورة في الرمية الأولى . حدد عناصر فراغ العينة ثم بين فيما إذا كانت مستقلة ثنائياً .

الحل :

عدد عناصر فراغ العينة = 4 وهي : $\Omega = \{HH, HT, TH, TT\}$

$A = \{ HH , TT \}$ و $B = \{ HH , HT , TH \}$ و $C = \{ HH , HT \}$

$\Leftarrow$ $P(A) = \frac{2}{4}$ و $P(B) = \frac{3}{4}$ و $P(C) = \frac{2}{4}$

وعليه يكون

$$P(A \cap B) = \frac{1}{4} \neq P(A)P(B) = \frac{1}{2} \times \frac{3}{4} = \frac{3}{8}$$

أي أن A و B حدثان غير مستقلان . أيضاً ،

$$P(A \cap B) = \frac{1}{4} = P(A)P(C) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$$

وبذلك يكون A و C حدثان مستقلان . وأخيراً فإن :

$$P(B \cap C)\frac{2}{4} \neq P(B)P(C) = \frac{3}{4} \times \frac{1}{2} = \frac{3}{8}$$

ومنها نستنتج أن B و C حدثان غير مستقلين .

مثال (42) : بفرض أن أسرة لها ثلاثة أطفال ، وبفرض تساوي احتمال ولادة الولد مع ولادة البنت ، حدد عناصر فراغ العينة ثم وضح فيما إذا كان الحدثان A و B مستقلان من عدمه .

الحل :

عدد نقاط (عناصر) فراغ العينة يساوي 8 هي :

$$\Omega = \{ bbb , bbg , bgb , gbb , ggb , gbg , bgg , ggg \}$$

حيث b تعني ولد و g تعني بنت .

فإذا كان الحدث A يمثل أسرة لها اولاد وبنات ، والحدث B يمثل أسرة لها ولد واحد على الأقل

فإن : $A = \{ bbg, bgb, gbb , ggb , gbg , bgg \}$

وإن $B = \{ bbb , bbg , bgb , gbb , ggb , gbg , bgg \}$

إذن $A \cap B = A \Rightarrow P(A \cap B) = P(A) = \frac{6}{8} = \frac{3}{4}$

$$P(A)P(B)=\frac{6}{8}\times\frac{7}{8}=\frac{21}{32}\neq P(A\cap B)=\frac{3}{4}$$

وعليه فإن :

إذن A و B حدثان غير مستقلين .

لقد أشرنا فيما سبق إلى أنه إذا كان A و B حدثين مستقلين ، فإن وقوع أو عدم وقوع الحدث A لا علاقة له بوقوع أو عدم وقوع الحدث B وعليه فإذا كان A و B يحققان التعريف الرياضي للأحداث المستقلة فإنه أيضاً يصح القول بأن الحدثين A و B' مستقلان ، وكذلك الحدثان A' و B مستقلان ، وأيضاً الحدثان A' و B' مستقلان . وسوف نصوغ ذلك في النظرية التالية :

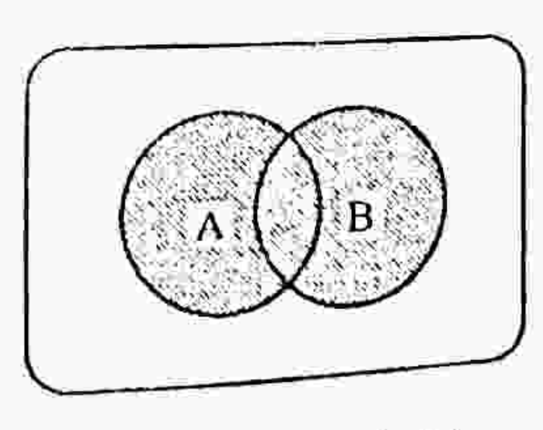
نظريـــة (10) : إذا كان A و B حدثين مستقلين معرفين على فراغ العينة Ω فإن (A,B') و (A',B) و (A',B') هي أزواج من حدثين مستقلين .

البرهان :

من الشكل (14) أدناه يتضح أن :

$$\begin{aligned}P(A\cap B') &= P(A)-P(A\cap B)\\ &= P(A)-P(A)P(B) \quad \text{لأن A و B مستقلان}\\ &= P(A)(1-P(B))\\ &= P(A)P(B')\end{aligned}$$

وعليه فإن الحدثان A و B' مستقلان .



شكل (15) : $A\cup B$

وبالمثل من الشكل (15) أعلاه نجد أن :

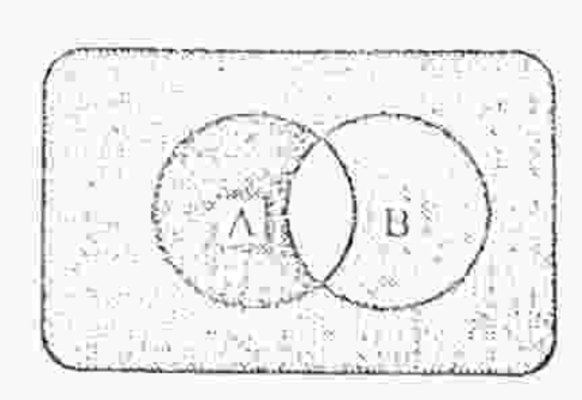
$$
\begin{aligned}
P(A' \cap B) &= P(B) - P(A \cap B) \\
&= P(B) - P(A)P(B) \\
&= P(B)(1 - P(A)) \\
&= P(B)P(A')
\end{aligned}
$$

وبالتالي فإن الحدثان A' و B مستقلان .

ومن الشكل (16) يتضح جلياً أن :

$$
\begin{aligned}
P(A' \cap B') &= P\left((A \cup B)'\right) \\
&= 1 - P(A \cup B) \\
&= 1 - \left[P(A) + P(B) - P(A \cap B)\right] \\
&= 1 - P(A) - P(B) + P(A \cap B) \\
&= 1 - P(A) - P(B) + P(A)P(B) \\
&= \left[1 - P(A)\right]\left[1 - P(B)\right] \\
&= P(A')P(B')
\end{aligned}
$$

وعليه فإن الحدثان A' و B' مستقلان .



شكل (16) : $A' \cap B' = (A \cap B)'$

ويمكن تعميم خاصية الاستقلالية لأكثر من حدثين . فإذا كانت A و B و C ثلاثة أحداث فإننا نقول أن A و B و C مستقلة ثنائيا (Pairwise Independent) إذا وفقط إذا (iff) تحققت الشروط الثلاثة التالية :

(I) $\quad P(A \cap B) = P(A)P(B)$

(II) $\quad P(A \cap C) = P(A)P(C)$

(III) $\quad P(B \cap C) = P(B)P(C)$

ولكن إذا تحقق الشرط التالي إضافة إلى الشروط الثلاثة أعلاه فإننا نقول أن الأحداث الثلاثة A و B و C مستقلة كلياً (Totally (or Mutually) Independent) :

(IV) $\quad P(A \cap B \cap C) = P(A)P(B)P(C)$

وإذا تحققت الشروط الأربعة سوية فإننا نقول أن الأحداث الثلاثة A و B و C مستقلة (Independent). ويمكن للقارئ أن يستنتج أنه إذا كان عدد الأحداث n فلابد من التحقق من شروط عددها $2^n - n - 1$ لإثبات الاستقلالية .

تعريف (22) : يقال بأن الأحداث $A_1, A_2, \ldots, A_n$ مستقلة عن بعضها البعض إذا تحققت الشروط الآتية :

(1) $\quad P(A_i \bigcap A_j) = P(A_i)P(A_j) \quad , 1 \le i \le j \le n$

(2) $\quad P(A_i \bigcap A_j \bigcap A_k) = P(A_i)P(A_j)P(A_k) \quad , 1 \le i < j < k < n$

(3) $\quad P(\bigcap_{i=1}^{n} A_i) = \prod_{i=1}^{n} P(A_i)$

وسنعرض الآن بعض الأمثلة التي توضح مفهوم وأهمية الاستقلالية في حل بعض المسائل في الاحتمالات .

مثال (43) : بفرض أن $\Omega = \{s_1, s_2, s_3, s_4\}$ و $P(s_i) = \frac{1}{4}$ حيث $i = 1 , 2 , 3 , 4$

وبفرض أن $A = \{s_1, s_2\}$, $B = \{s_1, s_3\}$, $C = \{s_1, s_4\}$ فإن :

$$A \cap B = A \cap C = B \cap C = \{s_1\} \quad , \quad A \cap B \cap C = \{s_1\}$$

وعليه فإن

$$P(A \cap B) = P(A \cap C) = P(B \cap C) = P(A \cap B \cap C) = \frac{1}{4}$$

$$P(A \cap B) = \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2} = P(A)P(B)$$

$$P(A \cap C) = \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2} = P(A)P(C)$$

$$P(B \cap C) = \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2} = P(B)P(C)$$

وبذلك تكون الأحداث A و B و C مستقلة ثنائياً . ولكن ، لما كان

$$P(A \cap B \cap C) = \frac{1}{4} \neq \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = P(A)P(B)P(C)$$

فإن الأحداث A و B و C غير مستقلة عن بعضها البعض .

مثال (44) : المطلوب إلقاء قطعة نقود متزنة تكراراً حتى الحصول على صورة لأول مرة، وبفرض أن نتائج الإلقاء مستقلة عن بعضها البعض ، ما احتمال أن تلقى قطعة النقود n مرة حتى يتحقق هذا الحدث ؟

الحل :

إن الاحتمال المطلوب الذي سنرمز له بالرمز P_n يساوي احتمال الحصول على n-1 كتابة متتالية ثم الحصول على صورة في الرمية n . وحيث أن نتائج إلقاء قطعة النقود مستقلة عن بعضها البعض فإن :

$$P_n = \left(\frac{1}{2}\right)^n$$

أما احتمال الحصول على صورة عاجلا أم آجلا ، أو عدم الحصول على كتابة على الإطلاق فهو

$$\sum_{n=1}^{\infty} P_n = \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \ \dots \ = 1$$

وحيث إن مجموع الاحتمالات يساوي 1 وعليه فإن احتمال الحصول على متوالية لا نهائية من الكتابات بدون الحصول على صورة على الإطلاق يجب أن يكون مساوياً للصفر .

مثال (45) : آلة تنتج سلعة معينة ، واحتمال إنتاجها وحدة معيبة يساوي p حيث أن $(0 < p < 1)$ ، واحتمال إنتاجها وحدة سليمة هو q حيث $q = 1-p$. وبفرض أنه تم اختيار عينة

تتكون من 6 وحدات من إنتاج تلك الالة بطريقة عشوائية ، وتم فحص هذه الوحدات فإذا كانت نتائج هذه العينة مستقلة عن بعضها البعض ، ما احتمال وجود وحدتين معيبتين في هذه العينة ؟

الحل :

لما كان فراغ العينة يحتوى على جميع الحالات الممكنة لهذه التجربة من حيث أن الوحدة قد تكون معيبة أو سليمة ، وعليه إذا كانت D تمثل حدث أن الوحدة معيبة (defective) و N حدث أنها سليمة (non - defective) ، وحيث أن نتائج الوحدات الستة مستقلة عن بعضها البعض ، فإن احتمال الحصول على أي متتالية معينة من الوحدات المعيبة والسليمة هو ببساطة حاصل ضرب الاحتمالات الفردية للوحدات . فعلى سبيل المثال :

$$\begin{aligned} P(NNDNDN) &= P(N)P(N)P(D)P(N)P(D)P(N) \\ &= q \cdot q \cdot p \cdot q \cdot p \cdot q \\ &= p^4 q^4 \end{aligned}$$

إذن يمكن ملاحظة أن احتمال أى متتالية معينة آخرى في Ω تحتوى على وحدتين معيبتين واربع وحدات سليمة ، سوف يكون ايضاً مساوياً إلى $p^2 q^4$. وعليه فإن احتمال وجود وحدتين معيبتين في عينة تتكون من ستة وحدات يمكن ايجاده بضرب الاحمال $p^2 q^4$ لاى متتالية معينة في عدد تلك المتواليات . وحيث إنه يوجد $\binom{6}{2}$ حالة ممكنة بها وحدتين معيبتين واربع وحدات سليمة فإن الاحتمال المطلوب هو :

$$\binom{6}{2} p^2 q^4$$

لنفرض أن المطلوب حساب احتمال وجود وحدة معيبة على الأقل من بين الوحدات الستة التى تم اختيارها . حيث أن احتمال أن تكون جميع الوحدات بالعينة سليمة هو q^6 فإن احتمال وجود وحدة معيبة على الاقل هو $1-q^6$.

مثال (46) : لنفرض أن كل علبة سجائر تحتوي على صورة لرئة مصابة نتيجة التدخين ، وبفرض أنه تستخدم r صورة مختلفة من هذه الصور لإظهار أضرار التدخين ، وأن لكل صورة نفس الفرصة بأن توضع في أي علبة وبطريقة مستقلة عن بقية الصور الأخرى ، ما احتمال أن شخصاً مدخناً اشترى n علبة من السجائر ($r \leq n$) سوف يتحصل على مجموعة الصور r المختلفة .

الحل :

لنفرض أن A_i حيث $i = 1 ، 2 ، . . . ، r$ يمثل أن الصورة i غير موجودة في جميع العلب التي عددها n . وبما أن الحدث $\bigcup_{i=1}^{r} A_i$ يمثل حدث الحصول على صورة واحدة على الأقل غير موجودة ، وعليه سوف توجد أولاً $P(\bigcup_{i=1}^{r} A_i)$.

حيث إن لأي صورة من الصور نفس الفرصة بأن توضع في أي علبة معينة ، وعليه فإن احتمال عدم وجود الصورة i في علبة معينة هو $\frac{r-1}{r}$ ، وبما أن العلب تتم تعبئتها بشكل مستقل فإن احتمال عدم وجود الصورة i في أي من العلب n هو $\left(\frac{r-1}{r}\right)^n$ وعليه فإن :

$$P(A_i)=\left(\frac{r-1}{r}\right)^n \qquad , i=1,2, \ \dots \ ,r$$

وبالمثل احتمال عدم وجود الصورتين i و j في جميع العلب هو :

$$P(A_i \cap A_j)=\left(\frac{r-2}{r}\right)^n$$

وبالمثل فإن احتمال عدم وجود ثلاث صور i و j و k في جميع العلب هو :

$$P(A_i \cap A_j \cap A_k)=\left(\frac{r-3}{r}\right)^n$$

وبالاستمرار بنفس الطريقة نجد أن احتمال عدم وجود جميع الصور في العلب هو :

$$P(A_1 \cap A_2 \cap A_3 \cap \cdots \cap A_r)=0$$

وعليه من النتيجة (1) بالبند (2 ـ 5) نستنتج أن :

$$P(\bigcup_{i=1}^{r} A_i)=r(\frac{r-1}{r})^n-\binom{r}{2}(\frac{r-2}{r})^n+\ \dots\ +(-1)^r\binom{r}{r-1}(\frac{1}{r})^n$$

$$=\sum_{j=1}^{r-1}(-1)^{j+1}\binom{r}{j}(1-\frac{j}{r})^n$$

وبما أن احتمال الحصول على مجموعة الصور (r) كاملة يساوي $1-P(\bigcup_{i=1}^{r} A_i)$ وعليه فإن :

$$P=\sum_{j=0}^{r-1}(-1)^{j}\binom{r}{j}(1-\frac{j}{r})^n$$

مما سبق يمكننا استنتاج أن قانون ضرب الاحتمالات في حالة الأحداث المستقلة وغير المستقلة يساعد في حساب الاحتمال $P(A \cap B)$ بطريقة غير مباشرة ، وذلك تفاديا لحسابه بطريقة مباشرة نتيجة لصعوبة ذلك في بعض الأحيان من ناحية ، ولعدم إمكانية حسابه من ناحية أخرى . فنحن نعلم مما سبق أنه إذا كان A و B حدثين مستقلين ومعرفين على نفس فراغ العينة فإن احتمال وقوعهما معاً هو عبارة عن حاصل ضرب احتمال حدوث كل منهما على حده أي أن :

$$P(A \cap B) = P(A)\,P(B)$$

بينما إذا كان A و B حدثين غير مستقلين فإن احتمال وقوع الحدثين معاً هو :

$$P(A \cap B) = P(A)\,P(B \mid A) \qquad , \quad P(A) > 0$$

$$= P(B)\,P(A \mid B) \qquad , \quad P(B) > 0$$

حيث تعتمد عملية الحساب على أي الحدثين وقع أولا . إضافة لما سبق إن قانون ضرب الاحتمالات يستخدم في تحديد ما إذا كانت الأحداث قيد الدراسة مستقلة أم لا . وخلاصة القول أن الاستقلالية لا تستخدم فقط في التعريف عما إذا كان الحدثان مستقلين أم لا ، ولكن تستخدم في وضع نموذج احتمالي لبعض التجارب .

تمرينات Exercises

1 - يحتوى صندوق على عشرون بطاقة منها عشرة حمراء مرقمة من 1 إلى 10 ، وعشرة بيضاء مرقمة من 1 إلى 10 ، وبفرض أنه سُحبت بطاقة من ذلك الصندوق ، فإذا كان الحدث A يمثل أن البطاقة المسحوبة تحمل رقم زوجي، والحدث B يمثّل أن البطاقة المسحوبة لونها بيضاء ، والحدث C يمثّل أن البطاقة المسحوبة تحمل رقم اقل من 5 . صف فراغ العينة ، ثم صف الأحداث الآتية لغوياً وكمجموعات جزئية من ذلك الفراغ .

أ) $A \cap B \cap C$ ب) $A \cap (B \cup C)$ جـ) $B \cap C'$

د) $A \cup B \cup C$ ه) $A' \cap B' \cap C'$

2 - إذا كان لدى لاعب رياضي ستة قمصان وأربعة أزواج من الجوارب فما هي عدد الطرائق التي يمكنه بها أن يرتدى القمصان والجوارب ؟

3 - إذا اشترك ثلاثة لاعبين من الفريق A وثلاثة لاعبين من الفريق B في مسابقة للعدو ، وعلمت بأن للاعبين الستة نفس الكفاءة فما احتمال أن المتسابقين من الفريق A سوف يفوزون بالترتيب الأول والثاني والثالث بينما المتسابقين من الفريق B سيفوزون بالترتيب الرابع والخامس والسادس ؟

4 - أوجد كل من :

أ) $\binom{90}{30}$ ، $\binom{50}{15}$ ، $\frac{(n+2)}{n!}$ ، $\frac{n!}{(n-1)}$ ، $7!$ ، $\frac{7!}{10!}$

ب) $\binom{n}{k}$ ، $\binom{n}{k-1}$ حيث n و k أعداد صحيحة موجبة و $k \leq n$.

5 - اثبت أن :

أ) $\sum_{i=0}^{n} \binom{n}{i} = 2^n$ ب) $\sum_{i=0}^{n} (-1)^i \binom{n}{i} = 0$

" إرشاد : أستخدم نظرية ذي الحدين "

6 - ما هي عدد الطرائق التي يمكن أن يجلس بها 3 طلاب وطالبتان في الحالات الآتية :

أ) في صف .

ب) في صف والطلبة بجوار بعضهم والطالبتان بجوار بعضهن البعض .

جـ) في صف ولكن الطالبتان بجوار بعضهن فقط .

7 - يحتوي صندوق على 10 كرات ، ما عدد الطرائق التي يمكن بها سحب أربعة كرات من هذا الصندوق في الحالات الآتية :

أ) بدون إعادة . ب) مع الإعادة .

8 - إذا كان على طالب الإجابة على 8 أسئلة من بين 10 أسئلة في أحد الامتحانات ، فأوجد :

أ) عدد الطرائق التي يمكن للطالب الإجابة بها على هذه الأسئلة .

ب) عدد الطرائق التي يمكن للطالب الإجابة بها على هذه الأسئلة إذا كان السؤالان الأول والثاني إجباريين .

جـ) عدد الطرائق التي يمكن للطالب الإجابة بها على الأسئلة إذا كان عليه الإجابة على سؤالين على الأقل من بين الأسئلة الأربعة الأولى .

9 - ما عدد الطرائق التي يمكن بها تقسيم 7 ألعاب على ثلاثة أطفال إذا علمت أن أصغرهم سيعطى ثلاثة ألعاب وتوزع البقية على الطفلين الآخرين بالتساوي .

10 - إذا جلس 10 أشخاص بطريقة عشوائية في صف يتضمن 20 مقعداً فما احتمال عدم جلوس أي اثنان منهم بجوار بعضهم ؟ وما احتمال جلوسهم على المقاعد التي بجوار بعضها البعض ؟

11 - إذا تم اختيار لجنة تتكون من أربعة أعضاء من بين 12 شخصاً فما احتمال اختيار شخصين معينين بهذه اللجنة ؟

12 - يحتوي صندوق على 20 مصباحاً كهربائياً من بينهم 4 مصابيح تالفة ، فإذا اختار شخص ثمانية مصابيح بطريقة عشوائية من هذا الصندوق واخذ شخص آخر بقية المصابيح فما احتمال أن تكون المصابيح الأربعة التالفة قد تم اختيارها من قبل نفس الشخص ؟

13 - إذا تم إلقاء ثمانية مكعبات نرد متزنة مرة واحدة ، فما احتمال أن كل رقم من الأرقام الستة سوف يظهر على الأقل مرة واحدة ؟

14 - يحتوي صندوق على 6 كرات حمراء وعلى عدد مجهول من الكرات البيضاء ، فإذا كان احتمال سحب كرتين ذات لون أحمر على التوالي ودون إعادة هو $\frac{1}{3}$ فما عدد الكرات البيضاء ؟

15 - يتكون فصل دراسي من 9 تلاميذ ، منهم اثنان اسماهما متشابهة وثلاثة آخرون أسمائهم متشابهة والأربعة الباقون أسمائهم أيضاً متشابهة ، فإذا جلس هولا التلاميذ على تسعة مقاعد في صف بطريقة عشوائية ، فما احتمال أن التلميذين اللذين اسماهما متشابهة سوف يجلسان على المقعدين الأول والثاني ، والثلاثة الذين أسمائهم متشابهة سيجلسون على المقاعد الثلاثة التي تلي الأول والثاني ، ويجلس الأربعة الذين أسمائهم متشابهة على المقاعد الأخرى الباقية ؟

16 - يحتوي صندوق على 25 بطاقة من بينها 12 بطاقة حمراء ، وبفرض أن هذه البطاقات سيتم توزيعها على ثلاث لاعبين A ، B ، C بطريقة عشوائية بحيث يستلم اللاعب A عشرة بطاقات واللاعب B ثمانية بطاقات ، واللاعب C سبعة بطاقات ، فما احتمال أن اللاعب A سوف يستلم ستة بطاقات حمراء واللاعب B بطاقتان حمراوان ، واللاعب C أربعة بطاقات حمراء ؟

17 - إذا كان احتمال نجاح طالب في أحد المقررات الدراسية هو 0.50 واحتمال نجاح طالب اخر في المقرر نفسه هو 0.20 واحتمال نجاح الطالبين هو 0.10 فأوجد :

أ) احتمال نجاح أحد الطالبين على الأقل .

ب) احتمال نجاح أحدهما فقط .

جـ) احتمال عدم نجاح أي منهما .

18 - إذا كان A و B حدثان حيث $P(A) = 1/3$ و $P(B) = 1/2$ فأوجد $P(B \cap A')$ في الحالات الآتية :

أ) A و B حدثان متنافيان . ب) $A \subset B$ جـ) $P(A \cap B) = 1/8$

19 - ألقي مكعبي نرد متزنين مرة واحدة، حدد عناصر فراغ العينة ثم أوجد الاحتمالات التالية:

أ) أن يكون مجموع الرقمين على الوجهين زوجي .

ب) أن يكون مجموع الرقمين على الوجهين أكبر من 6 .

جـ) أن يكون مجموع الرقمين على الوجهين فردى .

د) أن يكون الفرق بين الرقمين على الوجهين أكبر من 3 .

هـ) أن يكون الفرق المطلق بين الرقمين يساوى 4 .

و) الحصول على عدد فردى من أحد المكعبين وعدد زوجي من المكعب الآخر.

ز) الحصول على وجهين متشابهين .

20 - ألقيت قطعة نقود معدنية متزنة ثلاث مرات ، أكتب فراغ العينة ثم أوجد :

أ) احتمال عدم الحصول على صورة في الرميات الثلاثة .

ب) احتمال الحصول على صورة وكتابتين في الرميات الثلاثة .

جـ) احتمال الحصول على ثلاث وجوه متشابهة .

د) احتمال الحصول على صورة واحدة على الأقل .

هـ) احتمال الحصول على صورتين على الأكثر .

21 - ألقى مكعب نرد متزن وقطعتي نقود معدنيتين متزنتين مرة واحدة ، أكتب عناصر فراغ العينة ثم أوجد احتمال حدوث الأحداث التالية :

أ) الحصول على صورتين ورقم زوجي .

ب) الحصول على صورة وكتابة وعدد أكبر من 3 .

جـ) الحصول على كتابتين وعدد فردي .

22 . إذا كان 60 % من طلبة قسم الإحصاء مسجلين في مقرر رياضة I و 40 % مسجلين في مقرر مبادئ الحاسوب ، و 80 % مسجلين في مقرر لغة عربية I ، و 20 % مسجلين في مقرري رياضة I و مبادئ الحاسوب ، و10 % مسجلين في مقرري رياضة I ولغة عربية I و 20 % مسجلين في مقرري مبادئ الحاسوب ولغة عربية I ، و5 % مسجلين في المقررات الثلاثة . فما هي نسبة الطلبة المسجلين في :

أ) مقرر واحد على الأقل . ب) مقرر واحد فقط .

23 . يحتوي صندوق على 10 كرات حمراء ، و10 كرات بيضاء ، و10 كرات زرقاء ، إذا تم اختيار 5 كرات من هذا الصندوق بطريقة عشوائية وبدون إعادة ، فما احتمال أن لون واحد على الأقل لم يتم اختياره ؟

24 . إذا تم إلقاء ستة مكعبات نرد مرة واحدة ، فما احتمال أن تكون الأرقام الثلاثة متساوية ؟

25 . إذا كان A ، B ، C ثلاثة أحداث معرفة على نفس فراغ العينة ومستقلة ، حيث $P(A) = 1/4$ و $P(B) = 1/3$ و $P(C) = 1/2$ فأوجد :

أ) احتمال عدم حدوث أي منها . ب) احتمال حدوث واحد منها فقط .

26 . في مباريات كأس العالم لكرة السلة فريقين A و B سيلعبان متوالية من المقابلات مع بعضهما البعض ، وأول فريق يفوز في أربعة مقابلات سيكون الفائز بكأس العالم لكرة السلة ، فإذا كان احتمال فوز الفريق A في أي مقابلة مع الفريق B هو 1/3 . فما احتمال أن الفريق A سيفوز بالكأس ؟

27 . إذا كان A حدثاً ما بحيث $P(A) = 0$ وكان B أي حدث آخر ، اثبت أن A و B حدثين مستقلين .

28 . يحتوي صندوق على 10 كرات حمراء ، و10 كرات صفراء ، وبفرض أنه تم اختيار 5 كرات من الصندوق كرة في كل مرة ومع الإعادة . أوجد احتمال أن لون واحد على الأقل لم يتم اختياره من ضمن الكرات الخمسة .

29 - إذا كان A و B حدثان متنافيان حيث $P(B) > 0$ أوجد $P(A|B)$.

30 - بالرجوع إلى التمرين رقم (22) إذا تم اختيار طالب بطريقة عشوائية وكان مسجلاً على الأقل في مقرر واحد من المقررات الثلاثة ، فما احتمال أن يكون مسجل في مقرر رياضة I . وإذا كان الطالب مسجل في مقرر رياضة I فما احتمال أنه مسجل أيضاً في مقرر مبادئ الحاسوب ؟

31 - يحتوي صندوق على بطاقة بيضاء وأربعة بطاقات حمراء مرقمة كالآتي: A ، B ، C ، D فإذا تم اختيار بطاقتين من الصندوق بطريقة عشوائية وبدون إعادة فما احتمال أن تكون البطاقتان لونهما احمر في الحالات التالية :

أ) إذا علمت أن البطاقة A قد تم اختيارها .

ب) إذا علمت أنه على الأقل بطاقة حمراء قد تم اختيارها .

جـ) بدون معلومات أخرى .

32 - الجدول التالي يبين توزيع 100 شخص مصنفين حسب الجنس والحالة الاجتماعية :

الجنس / الحالة الاجتماعية	ذكور	إناث	المجموع
متزوج	20	26	46
أعزب	32	22	54
المجموع	52	48	100

إذا تم اختيار شخص بطريقة عشوائية فأوجد احتمال حدوث الأحداث التالية :

أ) أن يكون ذكر .

ب) أن يكون متزوج .

جـ) أن يكون متزوج وذكر .

د) أن يكون ذكر أو متزوج .

هـ) أن يكون متزوج إذا علمت انه أنثى .

و) أن يكون ذكر علما بأنه متزوج .

33 - إذا كان $P(A) > P(B)$ فاثبت أن $P(A|B) > P(B|A)$.

34 . إذا كان A و B حدثين معرفين على نفس فراغ العينة حيث $P(A) > 0$

و $P(A) + P(B) > 1$ فأثبت أن $P(B|A) \geq 1 - \dfrac{P(B')}{P(A)}$.

35 . إذا كان هناك أربعة طرق وهي A , B , C , D تؤدي إلى الهروب من أحد السجون ، وهرب سجين من هذا السجن واختار طريقه بطريقة عشوائية ، وعلمت أنه إذا اختار الطريق A فإن احتمال هروبه هو 1/8 وإذا اختار الطريق B فإن احتمال هروبه هو 1/6 ، وإذا أختار الطريق C فإن احتمال هروبه هو 1/4 أما إذا اختار الطريق D فإن احتمال هروبه هو 9/10 فأوجد

أ) احتمال أن السجين سوف ينجح في الهروب من السجن .

ب) احتمال أنه قد اختار الطريق D إذا علمت أنه نجح في الهروب من السجن .

جـ) احتمال أنه قد اختار الطريق A علماً بأنه نجح في الهروب من السجن .

36 . إذا كان الحدثان A و B معرفين على نفس فراغ العينة بحيث $P(A) = 0.3$

و $P(A \cup B) = 0.8$ و $P(B) = p$ فأوجد :

أ) قيمة p التي تجعل A و B حدثان مستقلان

ب) الشرط الضروري لكي يكون الحدثان A و B متنافيين .

37 . ألقي مكعبي نرد متزنين مرة واحدة ، فإذا كان الحدث A يمثل ظهور الأرقام 4 أو 5 أو 6 والحدث B يمثل ظهور الأرقام 1 أو 2 أو 3 ، والحدث C يمثل مجموع الرقمين على المكعبين يساوي 7 ، وضح فيما إذا كانت الأحداث A ، B ، C مستقلة ثنائياً ، مستقلة ، ولماذا ؟

38 . إذا علمت أن احتمال سقوط الأمطار على مدينة تاجوراء هو 0.50 واحتمال أن يكون الجو بارداً هو 0.70 ، واحتمال سقوط الأمطار بشرط أن يكون الجو بارداً هو 0.30 . ما احتمال أن يكون الجو بارداً أو تسقط الأمطار ؟ هل الحدثين مستقلين ؟

39 ـ إذا علمت أن احتمال تشابه الطقس (ممطر أو صحو) في يومين متتاليين هو 0.75 فاحسب ما يلي :

أ) احتمال أن يكون الجو صحو بعد غداً علما بأن الجو ممطر اليوم .

ب) احتمال أن يكون الجو ممطر بعد غداً علماً بأن الجو ممطر اليوم .

40 - إذا علمت أن 45 % من طلبة أحد المعاهد العليا ذكور و 55 % إناث ، وأن 50 %من الإناث و 40 % من الذكور مدخنين ، فإذا تم اختيار طالب بطريقة عشوائية ووجد أنه مدخن ما احتمال أن يكون ذكر ؟

41 - يحتوى صندوق على 12 نضيدة ، أربعة منها غير فاسدة ، فإذا تم اختيار نضيدتين بطريقة عشوائية أوجد احتمال :

أ) أن تكونا فاسدتين .

ب) أن تكونا صالحتين .

جـ) أن يكون إحداهما على الأقل فاسدة .

د) أن تكون أحدهما صالحة والأخرى فاسدة .

42 - أوجد الاحتمالات المطلوبة في تمرين (41) ، إذا تم سحب النضيدتين الواحدة تلو الأخرى وبدون إعادة .

43 - يصوب شخصان نحو هدف مشترك ، فإذا كان احتمال أن الشخص الأول يصيب الهدف هو 1/4 ، واحتمال أن الشخص الثاني يصيب الهدف هو 2/5 ، فما احتمال أن يصيب الهدف إحداهما على الأقل ؟

44 - يحتوي صندوق على 5 كرات حمراء و 6 كرات بيضاء ، ويحتوي صندوق آخر على 6 كرات بيضاء و 4 كرات حمراء ، فإذا اختيرت كرة من الصندوق الأول وبدون مشاهدة لونها ووضعت في الصندوق الثاني، ثم سحبت كرة من الصندوق الثاني فما احتمال أن تكون حمراء ؟

45 - يصوب ثلاثة أشخاص نحو هدف مشترك ، واحتمال أن يصيب كل منهم الهدف هو 1/3 ، 1/4 ، 1/6 على التوالي ، فإذا صوب كل منهم نحو الهدف مرة واحدة ، فما احتمال أن واحد منهم فقط سوف يصيب الهدف ؟ وإذا كان واحد منهم فقط أصاب الهدف ، فما احتمال أن يكون الشخص الأول ؟

الفصـــل الثـــالــــــــث

متغيرات عشوائية في بعد واحـــــــد

One - Dimension Random Variables

3 - 1 مقدمـــــة Introduction

من خلال دراستنا في الفصل الثاني يتضح أن فراغ (فضاء) العينة Ω الذي يتضمن جميع النتائج الممكنة للتجربة العشوائية قد يكون من الصعب كتابة عناصره ، وذلك لأن هذا الفراغ قد يكون محدودا وقد لا يكون محدود ، منفصلاً أو متصلاً وأيضا قد تكون عناصره أعداداً أو خلاف ذلك . علاوة على ذلك ما يهم الباحث في معظم التجارب هو نتائج عددية ، فمثلاً عند إلقاء قطعة عملة نقدية قد يهمنا معرفة عدد الصور أو الكتابات التي سوف تظهر عند رمى هذه القطعة عدداً من المرات ، وليس معرفة النتائج المؤلفة من متتابعة من الصور والكتابات ، أيضاً عند إلقاء زهرة (مكعب) نرد قد يهمنا معرفة فيما إذا سيكون مجموع الرقمين 6 وليس ما إذا كانت النتيجة هي (1،5) أو (4،2) أو (3،3) أو (2،4) أو (1،5) .

وعليه سوف نتعرض في هذا البند إلى الطريقة التي يمكن بها صياغة قاعدة أو مجموعة من القواعد التي تمكننا من تمثيل عناصر فراغ العينة Ω بأعداد ، ولنكن x ، علاوة على ذلك إن اهتمامنا لا يكون مقتصراً على عناصر فراغ العينة فقط بل على دوال في تلك العناصر . هذه الدوال سنطلق عليها تسمية متغير عشوائي .

فإذا فرضنا أن التجربة العشوائية تتمثل في إلقاء قطعة نقدية مرة واحدة وملاحظة وجهها العلوي فإن فراغ العينة المصاحب لهذه التجربة هو $\Omega = \{\omega : H \text{ أو } T\}$ حيث T و H يمثلان الكتابة والصورة على التوالي . ولنفرض أن X دالة بحيث أن :

$$X(\omega) = \begin{cases} 0 & ,\omega = T \\ 1 & ,\omega = H \end{cases}$$

أي أن $X(H) = 1$ و $X(T) = 0$.

وعليه فإن X دالة ذات قيمة حقيقية معرفة على فراغ العينة Ω وهذه الدالة تنقل الأحداث الموجودة في Ω إلى فراغ الأعداد الحقيقية $R_x = \{x : x = 0,1\}$ وفى مثل هذه الحالة نطلق على X تسمية متغير عشوائي والفراغ المصاحب له هو R_X . ويمكن توضيح هذه الحالة كما في شكل (1):



شكل (1): المتغير العشوائي X كدالة من Ω إلى R_X .

تعريف (1): إذا كانت Ω تمثل فراغ العينة لتجربة عشوائية ، فإن الدالة X ، التي تعطى عدد حقيقي $X(\omega)$ لكل ω في Ω تسمى متغيرًا عشوائيًا.

إذن يتضح من التعريف أن $X(\omega)$ تأخذ قيما على الخط الحقيقي R ، وفي الواقع سيكون هناك فراغا جديدًا وهو $R_X = \{x: X(\omega) = x, \omega \in \Omega\} \subseteq R$ والذي من الممكن أن يكون R .

وعادة ما يرمز للمتغير العشوائي بحرف كبير مثل W,U,V,Z,Y,X .... الخ . ولقيمة ذلك المتغير العشوائي بحرف صغير مثل w,u,v,z,y,x ...الخ .

مثال (1): إذا ألقيت قطعة عملة نقدية متزنة مرتين ، فإن Ω تتضمن أربعة نقاط ، فإذا فرضنا أن المتغير العشوائي X يمثل عدد الصور التي ستظهر في الرميتين فإن :

$$X(HH) = 2 \ , X(HT) = X(TH) = 1 \ , X(TT) = 0$$

أي أن X تأخذ القيم 2،1،0 . وفي هذه الحالة سوف يكون اهتمامنا متعلقاً بالأحداث المصاحبة بالفضاء $R_X = \{x: x = 0,1,2\}$ والمتغير العشوائي X سوف يحدث احتمالات على هذه الأحداث . ففي هذا المثال لكل عنصر من عناصر فضاء العينة نفس الفرصة في الحدوث أي أن احتمال حدوث أي منها يساوي $\frac{1}{4}$ ، فإذا كان الحدث A مثلاً يمثل ظهور صورة وكتابة فإن $A = \{HT, TH\}$ وهو مرتبط بالفضاء Ω بينما الحدث $\{X=1\}$ مرتبط بالفضاء R_X ولكن

$$P(X = 1) = \sum_{\omega} P(\omega: X(\omega) = 1) = P(A) = \frac{2}{4} = 0.5$$

وذلك لأن الحدث A حدث مكافئ (equivalent event) في Ω والاحتمال معرف على الأحداث التي بفضاء العينة ، وإن المتغير العشوائي X أخذ الاحتمال 0.5 للحدث $(X = 1)$ كما في شكل (2). وهكذا بالنسبة لبقية القيم التي من الممكن أن يأخذها المتغير العشوائي X .ألحظ أن الفراغ الجديد في هذا المثال هو $R_X \subset R$ وإن جميع الفئات الجزئية تمثل أحداثا يمكن حساب احتمالاتها أيضاً ، وبصفة عامة سوف نستخدم الرمز $P_X(X=x)$ أو ببساطة $P(X=x)$ عند حساب احتمال حدوث حدث في مدى المتغير العشوائي X .



شكل (2) : عدد الصور في الرميتين .

مثال (2): إذا ألقيت زهرة نرد متزنة مرة واحدة ، فإن $\Omega=\{1,2,3,4,5,6\}$ ، وبالتالي فإن النتيجة هنا عدد حقيقي ، وإذا فرضنا أن المتغير العشوائي X يمثل عدد النقاط التي ستظهر على الوجه العلوي فإن $X(\omega) = \omega$ وعليه فإن الحدث $(X=2)$ يحدث إذا وفقط إذا كانت $\omega \in \{1,2,3,4,5,6\}$ وإن $P(X = 2) = 1/6 \approx 0.167$.

وإذا افترضنا أن الحدث A يمثل ظهور رقم فردي فإن $A = \{1,3,5\}$ وإذا افترضنا أن المتغير العشوائي X يمثل عدد النقاط الفردية فإنه يمكن تعريف هذا المتغير كما يلي :-

$$X(\omega) = \begin{cases} 1 & ,\omega = 1,3,5 \\ 0 & ,\omega = 2,4,6 \end{cases}$$

وعليه فإن $R_X = \{x : x = 0,1\}$ ، وإن

$$P(X = 1) = \sum_{\omega} P(\omega : X(\omega) = 1) = P(A) = \frac{3}{6} = 0.5$$

وذلك لأن الحدث A حدث مكافئ في Ω والاحتمال معرف على الأحداث التي بفضاء العينة ، ولإن المتغير العشوائي X أحدث الاحتمال 0.5 للحدث $(X = 1)$.

إن المفاهيم السابقة يمكن صياغتها في التعريفين التاليين .

تعريف (2) : إذا كانت Ω تمثل فضاء عينة لتجربة عشوائية و R_X تمثل فضاء المتغير العشوائي X المعرف على Ω ، وكان A حدث معرف على Ω بينما B حدث معرف على R_X، فإنه يقال بأن A و B حدثان متكافئان إذا كان $A = \{\omega \in \Omega : X(\omega) \in B\} = X^{-1}(B)$.

إذن وفقاً لهذا التعريف أنه عندما يحدث الحدث A فإن الحدث B سوف يحدث أيضاً والعكس بالعكس وإن $X^{-1}(B)$ يرمز لصورة B تحت معكوس X .

تعريف (3) : إذا كان $A \subseteq \Omega$ و $B \subseteq R_X$ وكانت $X^{-1}(B) = A$ فإن احتمال حدوث الحدث B معرف كما يلي :

$$P(B) = P_X(B) = P(X^{-1}(B)) = P(A) .$$

إذن وفقاً لهذا التعريف نعطى الاحتمالات للأحداث التي في R_X بدلالة الاحتمالات المعرفة في Ω ، وفي المستقبل نتعامل مع طبيعة الدالة X وذلك لأن ما يهمنا هو القيم التي بفضاء المتغير العشوائي والاحتمالات المصاحبة لها . ألحظ أنه من الممكن أن لا تكون النتائج التي بفضاء العينة أعداداً حقيقية ولكن جميع عناصر مدى المتغير العشوائي سوف تكون أعداداً حقيقية .

مثال (3) : إذا ألقيت قطعة عملة نقدية متزنة ثلاث مرات وكان المتغير العشوائي X يمثل عدد الصور في الرميات الثلاث فإن الأحداث التي بفضاء هذا المتغير والأحداث المكافئة لها بفضاء العينة والاحتمالات المصاحبة لها تكون كالآتي :

بعض الأحداث في R_X	الأحداث المكافئة في Ω	الاحتمال
$X = 0$	$\{(T,T,T)\}$	0.125
$X = 1$	$\{(T,T,H),(T,H,T),(H,T,T)\}$	0.375
$X = 2$	$\{(H,H,T),(H,T,H),(T,H,H)\}$	0.375
$X = 3$	$\{(H,H,H)\}$	0.125

إذن وفقاً للمفاهيم السابقة فإن الصورة العكسية عندما $X = 0$ هي { (T,T,T) }، وإن الصورة العكسية عندما $X = 1$ هي { (T,T,H) , (T,H,T) , (H,T,T) } ، والصورة العكسية عندما $X = 2$ هي { (H,H,T) , (H,T,H) , (T,H,H) } وأخيراً الصورة العكسية عندما $X = 3$ هي { (H,H,H) } ، وبالتالي فإن احتمال وقوع الحدث { $X = x$ } مساوياً لاحتمال وقوع الصورة العكسية له ، أي أن

$$P(X = x) = P(X^{-1} (x)) = P(\omega : X(\omega) = x)$$

وعليه فإن

$$P(X = 0) = P(X^{-1}(0)) = P(\{(T,T,T)\}) = 0.125$$

$$P(X = 1) = P(X^{-1}(1)) = P(\{ (T,T,H) , (T,H,T) , (H,T,T) \}) = 0.375$$

$$P(X = 2) = P(X^{-1}(2)) = P(\{ (H,H,T) , (H,T,H) , (T,H,H) \}) = 0.375$$

$$P(X = 3) = P(X^{-1}(3)) = P(\{ (H,H,H) \}) = 0.125$$

مثال (4) : إذا وضعت نضيدة في التشغيل عند الزمن $t = 0$ وراقبتها حتى تتوقف عن العمل ، فإن فراغ العينة في هذه الحالة سيكون $\Omega = (0,\infty)$ ، وإن الأحداث التي قد تكون قيد دراسة هنا هي الفترات الجزئية من Ω ، فإذا كان المتغير العشوائي X يمثل الفترة الزمنية التي تستغرقها النضيدة حتى تتوقف عن العمل ، فإن $X(\omega) = \omega$ ، ولكن كيف يمكن إعطاء الاحتمالات للأحداث في مثل هذه الحالة ؟ هذا ما سنناقشه في بند قادم .

لقد تكونت للقارئ فكرة عن المتغير العشوائي من خلال الأمثلة السابقة وقد ألحظ أن فراغ العينة Ω في الأمثلة الثلاثة الأولى كان مجموعة قابلة للعد وكانت قيم المتغير العشوائي محدودة. بينما في المثال (4) (حيث $\Omega = (0,\infty)$) كان فراغاً غير محدود وغير قابل للعد . وبهذا يتضح أن قيم المتغيرات العشوائية المعرفة على Ω من الممكن أن تكون على الأكثر عدداً من القيم القابلة للعد أو عدد من القيم غير قابلة للعد .

وفي الخلاصة يمكننا القول بأن المتغير العشوائي هو دالة حقيقية معرفة على فراغ العينة في تجربة عشوائية ، وإنه من الحكمة نقل النتائج الأصلية $\omega \in \Omega$ ، متى كان ذلك ضرورياً إلى أعداد حقيقية، والتي بدورها ستساعدنا في استخدام الخواص المألوفة لنظام الأعداد الحقيقية وأثناء هذه النقلة سنتخلص من كل المعلومات التي لسنا بحاجة إليها وغالبا ما يؤدي ذلك إلى دراسة فراغ جزئي $(R_X \subseteq R)$ اصغر بكثير .وأخيراً وكما أشرنا سابقاً للدلالة على قيمة معينة

للمتغير العشوائي سوف نستخدم نفس الرمز ولكنه مصغر ، وبالتالي فإن $(X = x)$ و $(X < x)$ و $(X \leq x)$ جميعها أحداث في فضاء المتغير العشوائي X .

تعريف (4): إذا كان X متغيراً عشوائياً وكانت مجموعة قيمه الممكنة منتهية أو غير منتهية ولكنها قابلة للعد ، فإننا نطلق على X تسمية متغير عشوائي منفصل أو متقطع (Discrete Random Variable)

ومن الأمثلة على المتغير العشوائي المنفصل :

ا- عدد الأطفال في الأسرة.

ب- عدد الأهداف التي سيسجلها فريق كرة قدم في مباراة قادمة .

جـ- عدد الأخطاء المطبعية في صفحات كتاب ما .

تعريف (5) : إذا كان X متغيراً عشوائياً وكانت مجموعة قيمه الممكنة تمثل فترة أو تجمع (collection) من الفترات فإنه يطلق على X تسمية متغير عشوائي متصل أو مستمر (Continuous Random Variable) .

ومن الأمثلة على المتغير العشوائي المتصل :

ا- حجم الغازات المنبعثة من انفجار بركاني محتمل الوقوع.

ب- الفترة الزمنية التي يعمرها مصباح كهربائي .

جـ- الفترة الزمنية التي تستغرقها عملية جراحية .

د- الطول ، الوزن ، العمر (سنة، شهر ، يوم ، ...) .

3-2 دالة التوزيع التراكمي Cumulative Distribution Function (C.D.F)

في كثير من الأحيان قد يتطلب الأمر حساب وتوضيح الاحتمالات المناظرة لكون متغير عشوائي يساوي أو أقل من قيمة معينة (أو أكبر منها)، ففي مثل هذه الحالة يفضل وجود دالة تربط قيم المتغير العشوائي قيد الدراسة بالاحتمالات التراكمية لقيم محددة وهذا ما يتضمنه التعريف التالي :

تعريف(6) : دالة التوزيع التراكمي للمتغير العشوائي X هي دالة نطاقها الخط الحقيقي R ومداها الفترة المغلقة $[0,1]$ ، ويرمز لها بالرمز $F_X(.)$ ومعرفة كما يلي :

$$F_X\ (x) = P(X \le x) = P(\{\omega : -\infty < X(\omega) \le x\}) \quad , x \in R \qquad (1)$$

والسبب في أهمية دالة التوزيع التراكمي ،هو أنها محددة بالكامل بتوزيع X كما أنه يمكن استخدامها لإيجاد احتمالات الأحداث المعرفة بدلالة المتغير العشوائي ،ولها الخواص التالية :

أ - $0 \le F_X(x) \le 1$ و $-\infty < x < \infty$

ب - $F_X\ (-\infty) = \lim\limits_{x \to -\infty} F_X\ (x) = 0$

أي أن دالة التوزيع التراكمي للقيم الصغيرة جدًا تساوي صفرًا ، ونقصد بالقيم الصغيرة جدًا هنا تلك القيم التي تكون أصغر من أصغر قيمة معطاة في السؤال أو التطبيق .وإن

$. F_X\ (+\infty) = \lim\limits_{x \to +\infty} F_X\ (x) = 1$

أي أن دالة التوزيع التراكمي للقيم الكبيرة جدًا تساوي واحدًا . ونقصد بالقيم الكبيرة جدًا هنا تلك القيم التي تكون أكبر من أكبر قيمة معطاة في السؤال أو التطبيق .

جـ - دالة التوزيع التراكمي دالة غير متناقصة ، بمعنى أنه إذا كانت $x_1 < x_2$ فإن $. F_X(x_1) \le F_X(x_2)$

د- دالة التوزيع التراكمي دائما متصلة من اليمين بمعنى أنه لجميع قيم x ، $h > 0$ يكون

$$\lim_{h \to 0}[F_X(x+h) - F_X(x)] = 0$$

كما أشرنا سابقا إذا كانت دالة التوزيع التراكمي للمتغير العشوائي X معرفة ، فإن احتمال أن X سوف تقع في أي فترة معينة على الخط الحقيقي يمكن تحديده بدالة التوزيع التراكمي والجدول الآتي يبين صيغ لحساب احتمالات أحداث معينة باستخدام هذه الدالة .

الحدث	الاحتمال
$\{X \le a\}$	$F_X\ (a)$
$\{X < a\}$	$F_X\ (a) - P(X = a)$
$\{X > a\}$	$1 - F_X\ (a)$
$\{X \ge a\}$	$1 - F_X\ (a) + P(X = a)$
$\{a < X \le b\}$	$F_X\ (b) - F(a)$
$\{a < X < b\}$	$F_X\ (b) - F_X\ (a) - P(X = b)$
$\{a \le X \le b\}$	$F_X\ (b) - F_X\ (a) + P(X = a)$
$\{a < X \le b\}$	$F_X\ (b) - F_X\ (a) + P(X = a)$
$\{a \le X < b\}$	$F_X\ (b) - F_X\ (a) + P(X = a) - P(X = b)$

وسوف نوضح في البند القادم والذي يليه كيفية إيجاد الدالة التراكمية وذلك استناداً إلى نوع المتغير العشوائي من حيث كونه منفصلاً أو متصلاً ، وعليه من وجهة نظر احتمالية يمكن تحديد نوعين رئيسيين من الدوال الاحتمالية للمتغيرات العشوائية ، هما دالة كتلة الاحتمال ودالة كثافة الاحتمال .

3 - 3 دالة كتلة الاحتمال لمتغير عشوائي منفصل (p.f.)
Probability mass function for a discrete random variable

لقد أشرنا فيما سبق أن المتغير العشوائي (X) المنفصل يأخذ قيماً قد تكون منتهية أو غير منتهية مثل : $x_1, x_2, \ldots, x_j, \ldots$ فإذا كانت الاحتمالات المناظرة لهذه القيم هي $p_X(x_1), p_X(x_2), \ldots, p_X(x_j), \ldots$ على التوالي والتي سنكتبها لغرض توضيح الفكرة بالشكل التالي (يطلق على الجدول التالي تسمية جدول التوزيع الاحتمالي للمتغير X) :

$X = x$	x_1	x_2	x_3	$\cdots$	$x_j \cdots$
$p_X(x) = P(X = x)$	$p_X(x_1)$	$p_X(x_2)$	$p_X(x_3)$	$\cdots$	$p_X(x_j)\ \cdots$

حيث أن $p_X(x_j) = P(X = x_j)$ لجميع $j = 1, 2, \cdots, n, \cdots$ فإننا نطلق على الدالة $p_X(x_j)$ المعرفة بالصيغة :

$$p_X(x_j) = \begin{cases} P(X = x_j) & , j = 1, 2, \cdots, n, \cdots \\ 0 & , \text{ow.} \end{cases} \qquad (2)$$

تسمية دالة كتلة احتمال X إذا وفقط إذا تحقق الشرطان التاليان :-

1- $0 \le p_X(x_j) \le 1$ لجميع قيم j .

2- $\sum_j p_X(x_j) = 1$

عليه إذا كان X متغيرًا عشوائيًا منفصلاً ، فإن احتمال حدوث أي مجموعة جزئية A من الخط الحقيقي يمكن تحديده بالعلاقة الآتية :

$$P(X \in A) = \sum_{x_j \in A} p_X(x_j) = \sum_{x_j \in A} P(X = x_j)$$
$$= \sum_{x_j \in A} P(\omega : X(\omega) = x_j)$$

وعليه فإن دالة كتلة الاحتمال (p.m.f.) للمتغير العشوائي المنفصل X هى دالة حقيقية نطاقها الخط الحقيقي ومداها الفترة [0,1] .وإن دالة التوزيع التراكمي (c.d.f.) للمتغير العشوائي X المنفصل تكون معرفة على النحو التالى :-

$$F_X(x) = P(X \le x) = \sum_{x_j \le x} p_X(x_j) \qquad (3)$$

وتفي بجميع الخواص السالف ذكرها في التعريف العام لدالة التوزيع التراكمي .

نظرية (1) : إذا كان X متغيراً عشوائياً منفصلاً فإنه يمكن الحصول على دالة التوزيع التراكمي $F_X(.)$ من دالة كتلة الاحتمال $p_X(.)$ والعكس صحيح .

البرهان :

بفرض أن $x_1, x_2, \ldots, x_k, \ldots$ تمثل قيم للمتغير العشوائي X وبفرض أن $p_X(.)$ معطاة إذن

$$F_X(x) = P(X \le x) = \sum_{\{j : x_j \le x\}} p_X(x_j)$$

الآن بفرض أن

$$x_1 < x_2 < \ldots < x_k < \ldots$$

إذن

$$F_X(x_j) = P(X \le x_j) = \sum_{i=1}^{j} p_X(x_i)$$

وايضا

$$F_X(x_{j-1}) = P(X \le x_{j-1}) = \sum_{i=1}^{j-1} p_X(x_i)$$

$$\Rightarrow F_X(x_j) - F_X(x_{j-1}) = \sum_{i=1}^{j} p_X(x_i) - \sum_{i=1}^{j-1} p_X(x_i) = p_X(x_j) \quad , j=1,2,\ldots,k,\ldots$$

ولتوضيح ما سبق ندرس الأمثلة الآتية .

مثال (5) : إذا كان المطلوب اختيار طالبين بطريقة عشوائية من بين 3 طلاب و3 طالبات وكان المتغير العشوائي X يمثل عدد الطالبات اللواتي سيتم اختيارهن ، فأوجد التوزيع الاحتمالي لهذا المتغير .

الحـل :

عدد الطرائق التي يمكن بها اختيار طالبين من بين 6 يساوي $\binom{6}{2} = 15$ طريقة ، وعليه فإن فضاء العينة يتضمن 15 نقطة ولكل عنصر نفس الفرصة في الظهور وذلك لأن المعاينة عشوائية ، وإن القيم الممكنة لهذا المتغير هي 0 ، 1 ، 2 ، وإن عدد الطرائق التي يمكن بها اختيار طالبة وطالب مثلاً أي أن $X = 1$ يساوي $9 = \binom{3}{1}\binom{3}{1}$ وهكذا لبقية القيم الممكنة ، وعليه فإن

$$p_X(0) = P(X=0) = \frac{\binom{3}{0}\binom{3}{2}}{15} = \frac{3}{15} = \frac{1}{5}$$

$$p_X(1) = P(X=1) = \frac{\binom{3}{1}\binom{3}{1}}{15} = \frac{9}{15} = \frac{3}{5}$$

$$p_X(2) = P(X=2) = \frac{\binom{3}{2}\binom{3}{0}}{15} = \frac{3}{15} = \frac{1}{5}$$

إن أفضل طريقة لتمثيل التوزيع الاحتمالي المنفصل هو كتابته في صيغة رياضية ويمكن كتابة دالة كتلة الاحتمال لهذا المثال كما يلي :

$$p_X(x) = P(X = x) = \begin{cases} \dfrac{\binom{3}{x}\binom{3}{2-x}}{\binom{6}{2}} & , \quad x = 0,1,2 \\ 0 & , \quad ow. \end{cases}$$

مثال (6) : ألقيت قطعة عملة نقدية متزنة حتى ظهور صورة لأول مرة ، فإذا كان المتغير العشوائي X يمثل عدد المحاولات المطلوبة لإلقاء العملة حتى ظهور الصورة فأوجد التوزيع الاحتمالي لهذا المتغير العشوائي X .

الحـل :

إذا كانت H ترمز للصورة و T ترمز للكتابة فإن النتائج الممكنة لهذه التجربة والاحتمالات المناظرة لها تكون كالآتي :

الرمية رقم	النتائج الممكنة	الاحتمال
1	H	$\frac{1}{2}$
2	TH	$(\frac{1}{2})(\frac{1}{2})$
3	TTH	$(\frac{1}{2})(\frac{1}{2})(\frac{1}{2})$
⋮		
x	TT…TH	$(\frac{1}{2})^{x-1}(\frac{1}{2})$

عليه فإن التوزيع الاحتمالي للمتغير العشوائي يمكن كتابته كما يلي :

$X = x$	1	2	…	x
$p_X(x) = P(X = x)$	$\frac{1}{2}$	$(\frac{1}{2})^2$	…	$(\frac{1}{2})^x$

بصيغة أخرى يمكن كتابة دالة كتلة الاحتمال كما يلي

$$p_X(x) = (\tfrac{1}{2})^x , \quad x = 1, 2, \ldots$$

ولتوضيح أن مجموع الاحتمالات يساوى واحد يجب جمع الاحتمالات وذلك كما يلي :

$$\sum_{x=1}^{\infty}\left(\tfrac{1}{2}\right)^{x}=\tfrac{1}{2}+\left(\tfrac{1}{2}\right)^{2}+\left(\tfrac{1}{2}\right)^{3}+\cdots$$

وهى تمثل سلسلة هندسية ولهذه السلسة يكون

$$a+ar+ar^{2}+\cdots=\tfrac{a}{1-r}$$

حيث a تمثل الحد الأول و r تمثل التناسب ما بين كل حدين متتالين ، وفي هذه الحالة $a=\frac{1}{2}$ و $r=\frac{1}{2}$ وعليه فإن

$$\sum_{x=1}^{\infty}p_{X}(x)=\frac{\frac{1}{2}}{1-\frac{1}{2}}=1.$$

مثال (7) : إذا ألقيت زهرة نرد متزنة مرة واحدة وكان المتغير العشوائي X يمثل عدد النقاط التي ستظهر على الوجه العلوي للزهرة . فأوجد

أ- دالة كتلة احتمال المتغير العشوائي X ومثلها بيانياً .

ب- دالة التوزيع التراكمي ومثلها بيانيا .

الحل :

أ- إن فراغ العينة فى هذه الحالة هو

$$\Omega=\{1,2,3,4,5,6\}$$

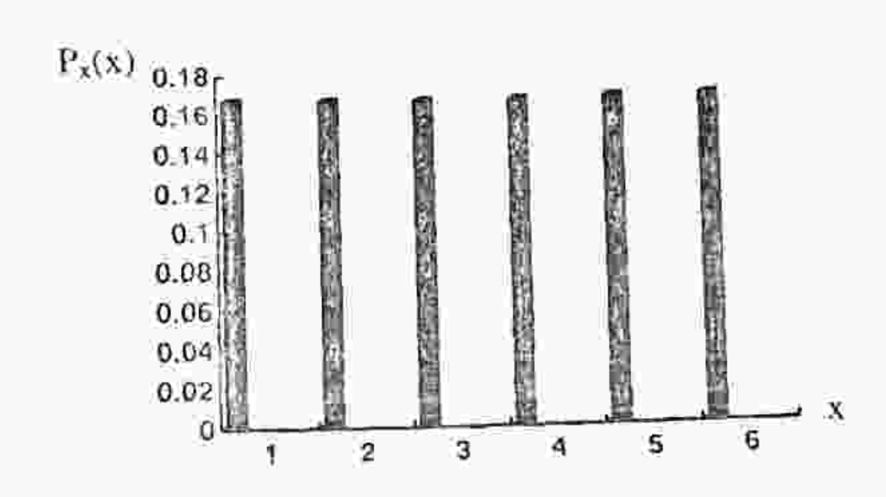
حيث أن المتغير العشواني X يمثل عدد النقاط التي ستظهر على الوجه العلوي فإن القيم الممكنة لهذا المتغير هى الأعداد الصحيحة من1 الى 6 . وحيث أن احتمال ظهور أى منها يساوى $\frac{1}{6}$ (لأن الزهرة متزنة) وعليه فإن دالة كتلة الاحتمال لهذا المتغير العشوائي يمكن كتابتها على النحو التالي :

$$p_{X}(x)=P(X=x)=\begin{cases}\frac{1}{6} & ,x=1,2,3,4,5,6\\ 0 & ,ow.\end{cases}$$

وقد يكون من المناسب في بعض الأحيان وضع دالة كتلة الاحتمال أعلاه في جدول على النحو الأتي :

$X=x$	1 2 3 4 5 6	$\sum_{x=1}^{6} P(X=x)$
$p_x(x)=P(X=x)$	$\frac{1}{6}$ $\frac{1}{6}$ $\frac{1}{6}$ $\frac{1}{6}$ $\frac{1}{6}$ $\frac{1}{6}$	$\frac{6}{6}=1$

واضح أن $P(X=x)$ تمثل دالة كتلة احتمال حيث أنها دالة غير سالبة وإن مجموع الاحتمالات المقترنة بقيم المتغير العشوائي X تساوى واحد ، ويمكن تمثيلها بيانياً كما يلي :

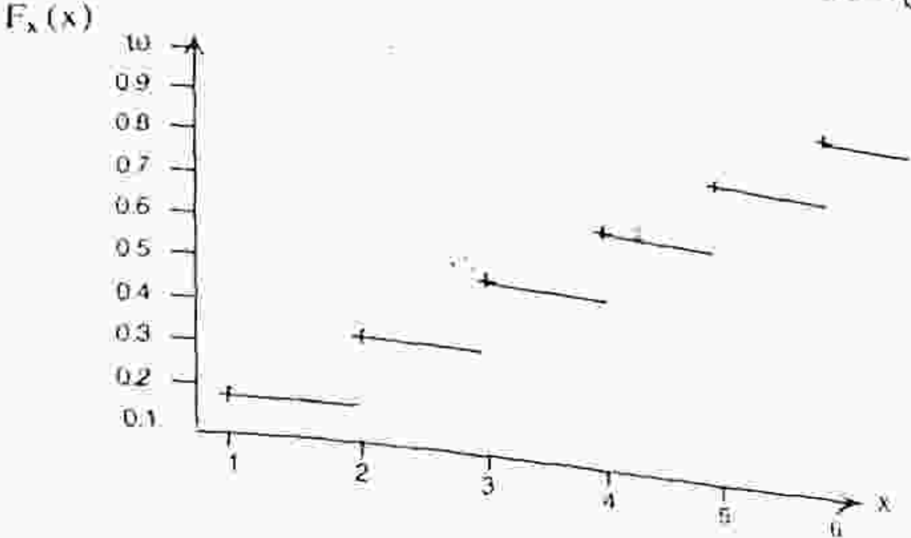


ب – دالة التوزيع التراكمي :

من دالة كتلة الاحتمال يمكن الحصول على دالة التوزيع التراكمي للمتغير العشوائي X على النحو التالي :

$$F_X(x)=P(X\le x)=\sum_{\{j:x_j\le x\}} P(X=x_j)=\begin{cases}0 & ,x<1\\ \frac{1}{6} & ,1\le x<2\\ \frac{2}{6} & ,2\le x<3\\ \frac{3}{6} & ,3\le x<4\\ \frac{4}{6} & ,4\le x<5\\ \frac{5}{6} & ,5\le x<6\\ 1 & ,x\ge 6\end{cases}$$

والتي يمكن تمثيلها بيانياً كالآتي :



شكل (4) :دالة التوزيع التراكمي للمتغير العشوائي X بمثال 7 .

ومن دالة التوزيع التراكمي يمكن حساب الاحتمال التراكمي لغاية أى قيمة من قيم X المعرفة في Ω وذلك بمجرد التعويض عن تلك القيمة في الدالة $F_X(x)$. فمثلاً

$$F_X(3) = P(X \leq 3) = \frac{3}{6} = 0.5$$

وأيضا يمكن حساب

$$p_X(3) = P(X = 3) = F_X(3) - F_X(2) = \frac{3}{6} - \frac{2}{6} = \frac{1}{6} = 0.167$$

أو مباشرة من دالة كتلة الاحتمال ،حيث

$$p_X(3) = P(X = 3) = \frac{1}{6} = 0.167 \ .$$

مثال (8) : إذا كان X متغيراً عشوائياً بدالة كتلة احتمال معرفة كما يلي :

$$p_X(x) = P(X = x) = \begin{cases} (\frac{1}{4})(\frac{3}{4})^x & ,x = 0,1,2... \\ 0 & ,ow. \end{cases}$$

فأوجد دالة التوزيع التراكمي ومنها أوجد $P(X \leq 3)$ و $P(2 < X \leq 5)$ و $P(X = 4)$.

الحـل :

من التعريف دالة التوزيع التراكمي تكون كالآتي :

$$F_X(x) = P(X \le x) = \sum_{y=0}^{x} (\frac{1}{4})(\frac{3}{4})^y$$

$$= \frac{1}{4}\sum_{y=0}^{x}(\frac{3}{4})^y = \frac{1}{4}[\frac{1-(\frac{3}{4})^{x+1}}{1-\frac{3}{4}}]$$

$$= 1-(\frac{3}{4})^{x+1} \qquad , x = 0,1,2,\dots$$

وبمعلومية دالة التوزيع التراكمي يمكن حساب الاحتمالات التالية :

$$P(X \le x) = F_X(3) = 1-(\frac{3}{4})^4 \cong 0.68$$

$$P(2 < X \le 5) = F_X(5) - F_X(2) = (\frac{3}{4})^3 - (\frac{3}{4})^6 \cong 0.38$$

$$P(X = 4) = F_X(4) - F_X(3) = (\frac{3}{4})^4 - (\frac{3}{4})^5 \approx 0.079$$

أو

$$p_X(4) = P(X = 4) = \left(\frac{1}{4}\right)\left(\frac{3}{4}\right)^4 = 0.079$$

مثال (9): يفرض أن X متغير عشوائي بدالة كتلة احتمال معطاه كالآتي :

$$p_X(x) = P(X = x) = \begin{cases} c\binom{4}{x} & , x = 0,1,2,3,4 \\ 0 & , ow \end{cases}$$

حيث c مقدار ثابت .أوجد كلاً من

أ- قيمة الثابت c ثم مثل الدالة بيانياً.

ب- دالة التوزيع التراكمي ومثلها بيانياً .

الحـل :

بما أن $p_X(x)$ دالة كتلة احتمال ، وعليه فإن

$$\sum_{x=0}^{4} p_X(x) = 1$$

$$\Rightarrow \sum_{x=0}^{4} c\binom{4}{x} = c\sum_{x=0}^{4}\binom{4}{x} = 1$$

$$c[\binom{4}{0} + \binom{4}{1} + \binom{4}{2} + \binom{4}{3} + \binom{4}{4}] = 1$$

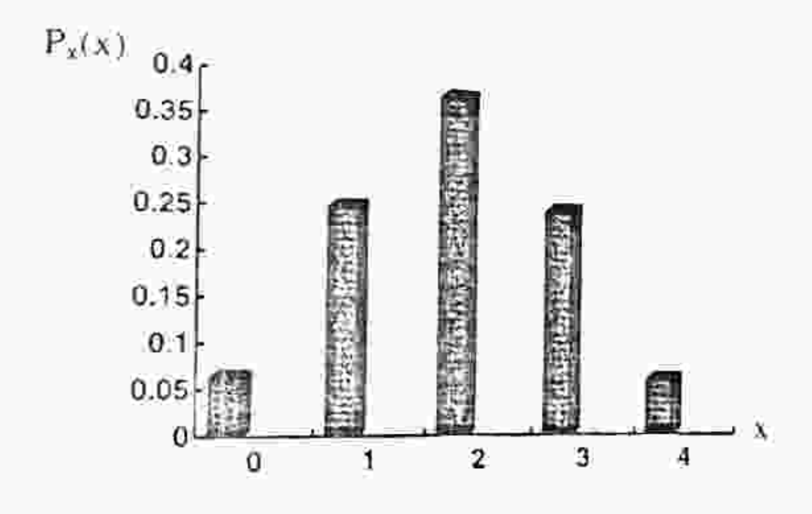
$$\Rightarrow \qquad 16c = 1$$

ومنها نجد أن $c = \frac{1}{16}$.

وبذلك فإن دالة كتلة الاحتمال للمتغير العشوائي X تكون على الصورة الآتية :

$$p_X(x) = P(X = x) = \begin{cases} \frac{\binom{4}{x}}{16} & , x = 0,1,2,3,4 \\ 0 & , ow. \end{cases}$$

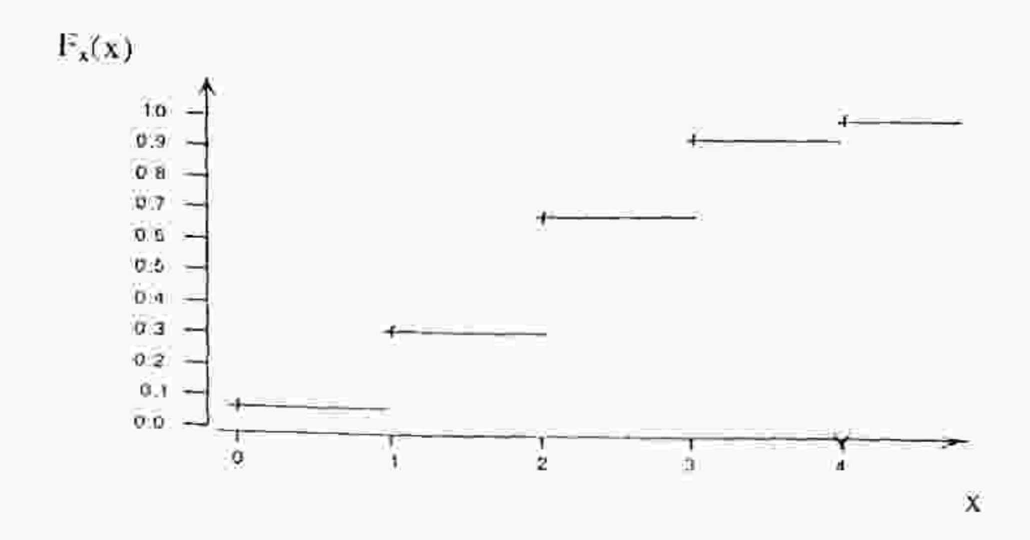
ويمكن تمثيلها بيانيا كما في الشكل أدناه:.



شكل (5) : دالة كتلة الاحتمال للمتغير العشوائي X بمثال 9 .

ب- دالة التوزيع التراكمي للمتغير العشوائي X يمكن إيجادها كما يلي :

$$F_X(x) = P(X \le x) = \begin{cases} 0 & ,x < 0 \\ \frac{1}{16} & ,0 \le x < 1 \\ \frac{5}{16} & ,1 \le x < 2 \\ \frac{11}{16} & ,2 \le x < 3 \\ \frac{15}{16} & ,3 \le x < 4 \\ 1 & ,x \ge 4 \end{cases}$$



شكل (6): دالة التوزيع التراكمي للمتغير العشوائي X بمثال 9 .

من الأمثلة السابقة يمكن أن نستنتج أنه إذا كان X متغيرا عشوائيا منفصلاً فإن دالة توزيعه التراكمية تتزايد على شكل قفزات فقط، وتكون ثابتة في الفترات التي لا تحصل فيها القفزات . وتحصل قفزات دالة التوزيع التراكمي هذه عند قيم المتغير العشوائي المنفصل X ويكون ارتفاع القفزة هو الاحتمال المناظر .

3 - 4 دالة كثافة الاحتمال لمتغير عشوائي متصل

The Probability Density Function of a Continuos Random Variable

يقال إن للمتغير العشوائي X توزيعاً متصلاً(مستمراً) إذا وجدت دالة غير سالبة $f_X(.)$ معرفة على الخط الحقيقي R ،بحيث أنه لأي فترة $A \subseteq R$ يكون :

$$P(X \in A) = \int_A f_X(x)\,dx \qquad ,A \subseteq R \qquad (4)$$

إن الدالة $f_X(.)$ تسمى دالة كثافة الاحتمال للمتغير العشوائي X .

إذن إذا كان X متغيرا عشوائيا بتوزيع متصل عندئذ يمكن إيجاد احتمال أن تنتمي X إلى أي مجموعة جزئية من الخط الحقيقي وذلك بإجراء التكامل لدالة كثافة احتمال المتغير X على تلك المجموعة الجزئية . ولابد لكل دالة كثافة احتمال أن تفي بالشرطين التاليين :

1- $f_X(x) \geq 0 \quad , \forall\ x \in R$

2- $\int_{-\infty}^{\infty} f_X(x) = P(R) = P(\Omega) = 1$

وإن دالة التوزيع التراكمي لهذا المتغير تكون متصلة على الخط الحقيقي بالكامل وتعرف كما يلي:

$$F_X(x) = P(X \leq x) = \int_{-\infty}^{x} f_X(u)\,du \qquad (5)$$

ولهذه الدالة جميع الخواص التي تم التطرق إليها في التعريف العام لدالة التوزيع التراكمي. وعند أي نقطة x تكون عندها دالة كثافة الاحتمال $f_X(.)$ متصلة فإن دالة التوزيع قابلة للتفاضل وإن

$$\frac{dF_X(x)}{dx} = f_X(x) \quad \Rightarrow dF_X(x) = f_X(x)\,dx$$

وهذا غالبا ما يطلق عليه " التفاضل الاحتمالي " للمتغير X .ومنها نستنتج أنه بمعلومية دالة التوزيع يمكننا أيجاد دالة كثافة الاحتمال والعكس صحيح أيضا، ويمكن استخدام دالة كثافة الاحتمال لحساب احتمال وقوع X في فترة معينة مثل $(a, b]$ وذلك كالآتي :

$$P(a \leq X \leq b) = \int_{a}^{b} f_X(x)\,dx \quad ,a < b$$

أو بدلالة دالة التوزيع وذلك على النحو الآتي :

$$P(a \leq X \leq b) = \int_{-\infty}^{b} f_X(x)\,dx - \int_{-\infty}^{a} f_X(x)\,dx = F_X(b) - F_X(a)$$

ونود أن نشير هنا إلى أنه إذا كان X متغيرا عشوائيا متصلاً فإن $P(\ X = x\ ) = 0$ ونتيجة لذلك فإنه لجميع قيم a و b بحيث أن $b > a$ يكون :

$$P(a < X \leq b) = P(a \leq X < b)$$
$$= P(a \leq X \leq b)$$
$$= P(a < X < b)$$
$$= \int_a^b f_X(x)\, dx$$

مثال (10) : بفرض أن X متغير عشوائي بدالة كثافة احتمال معرفة كما يلي :

$$f_X(x) = \begin{cases} \frac{1}{2} & , \; 5 \leq x \leq 7 \\ 0 & , \text{ow.} \end{cases}$$

أوجد ما يلي :

أ - $P(5.75 \leq X \leq 6.25)$ ب - $P(X \leq 6)$ جـ - $P(X \leq t)$ حيث t ما بين 5 و 7 .

الحـل :

أ -

$$P(5.75 \leq X \leq 6.25) = \int_{5.75}^{6.25} \left(\tfrac{1}{2}\right) d\,x = \left[\tfrac{1}{2}\,x\right]_{5.75}^{6.25}$$
$$= \tfrac{1}{2}(6.25 - 5.75) = 0.25.$$

ب -

$$P(X \leq 6) = P[-\infty < X \leq 6] = \int_{-\infty}^{6} f_X(x)\, d\,x$$
$$= \int_{-\infty}^{5} 0\, d\,x + \int_{5}^{6} \tfrac{1}{2}\, d\,x$$
$$= 0 + \left[\tfrac{1}{2}\,x\right]_5^6 = \tfrac{1}{2}(6-5) = \tfrac{1}{2}.$$

جـ -

$$P(X \leq t) = P[-\infty < X \leq t] = \int_{-\infty}^{t} f(x)\, d\,x$$
$$= \int_{-\infty}^{5} 0\, d\,x + \int_{5}^{t} \tfrac{1}{2}\, d\,x$$
$$= 0 + \left[\tfrac{1}{2}\,x\right]_5^t = \frac{t-5}{2}$$

مثال (11) : إذا كان X متغيراً عشوائياً بدالة كثافة احتمال معرفة كما يلي :

$$f_X(x) = \begin{cases} ax & ,0 \le x \le 1 \\ a & ,1 \le x \le 2 \\ -ax + 3a & ,2 \le x \le 3 \\ 0 & ,ow. \end{cases}$$

حيث a مقدار ثابت .أوجد قيمة الثابت a ثم أوجد $P(X \le 1.5)$.

الحـل :

من التعريف لكي تكون الدالة $f_X(x)$ دالة كثافة احتمال يجب أن تحقق الشرط :

$$\int_{-\infty}^{\infty} f_X(x)\,dx = 1$$

وعليه فإن :

$$\int_{-\infty}^{0} f_X(x)\,dx + \int_{0}^{1} f_X(x)\,dx + \int_{1}^{2} f_X(x)\,dx + \int_{2}^{3} f_X(x)\,dx + \int_{3}^{\infty} f_X(x)\,dx = 1$$

$$\Rightarrow \int_{0}^{1} ax\,dx + \int_{1}^{2} a\,dx + \int_{2}^{3} (-ax + 3a)\,dx = 1$$

$$\Rightarrow \frac{a}{2} + a + \frac{a}{2} = 1 \quad \Rightarrow \quad 2a = 1 \quad \Rightarrow \quad a = 0.5$$

وعلى فإن دالة كثافة الاحتمال (p . d . f.) للمتغير العشوائي X تكون كالاتي :

$$f_X(x) = \begin{cases} 0.5x & ,0 \le x \le 1 \\ 0.5 & ,1 \le x \le 2 \\ -0.5x + 1.5 & ,2 \le x \le 3 \\ 0 & ,ow. \end{cases}$$

وإن الاحتمال المطلوب يتم حسابه كما يلي :

$$P(X \le 1.5) = \int_{-\infty}^{1.5} f_X(x)\,dx = \int_{0}^{1} f_X(x)\,dx + \int_{1}^{1.5} f_X(x)\,dx$$

$$= \int_{0}^{1} 0.5x\,dx + \int_{1}^{1.5} 0.5\,dx = 0.25 + 0.25 = 0.5.$$

مثال (12) : بفرض أن X متغير عشوائي بدالة معرفة كما يلي :

$$f_X(x)=\begin{cases} x & ,0<x\le 1 \\ 2-x & ,1\le x\le 2 \\ 0 & ,ow. \end{cases}$$

والمطلوب :

أ- اثبت أن الدالة دالة احتمالية ومثلها بيانياً .

ب- أوجد دالة التوزيع التراكمي ومثلها بيانياً .

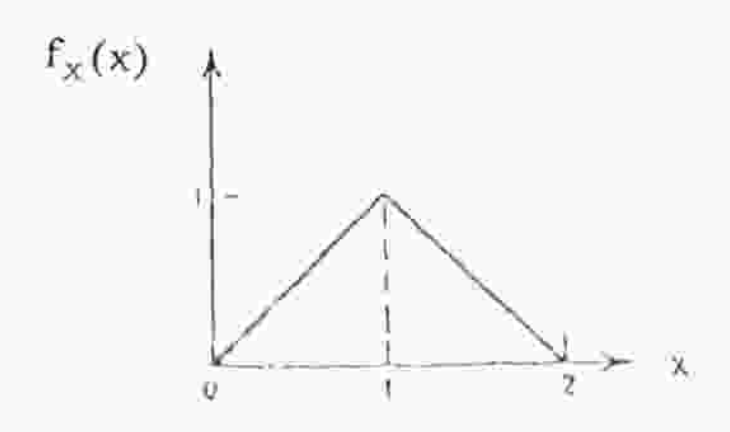
جـ - أوجد $P(0.3<X\le 1.5)$.

الحـل :

أ- حيث أن $f_X(x)\ge 0$ لجميع قيم x ، وكذلك

$$\int_{-\infty}^{\infty} f_X(x)\,dx=\int_0^1 x\,dx+\int_1^2 (2-x)\,dx=\frac{x^2}{2}\bigg|_0^1+(2x-\frac{x^2}{2})\bigg|_1^2=1$$

وبالتالي فإن $f_X(x)$ دالة كثافة احتمالية وشكلها يظهر على النحو التالي :



شكل (7) : دالة كثافة الاحتمال $f_X(x)$ للمتغير العشوائي **X في مثال (12)** .

ب- دالة التوزيع التراكمي $F_X(x)$ للمتغير العشوائي X ستكون كالآتي :

(1) عندما تكون $x\le 0$ فإن :

$$F_X(x)=\int_{-\infty}^{x} 0\,dy=0$$

(ii) عندما تكون $0 < x \leq 1$ فإن :

$$F_X(x) = \int_{-\infty}^{x} f_X(y)\,dy = \int_{-\infty}^{0} 0\,dy + \int_{0}^{x} y\,dy = 0 + \frac{x}{2} = \frac{x}{2}$$

(iii) عندما تكون $1 < x \leq 2$ فإن :

$$F_X(x) = \int_{-\infty}^{x} f_X(y)\,dy = \int_{-\infty}^{0} 0\,dy + \int_{0}^{1} y\,dy + \int_{1}^{x} (2-x)\,dy$$

$$= 2x - \frac{x^2}{2} - 1$$

(iv) عندما $x \geq 2$ فإن :

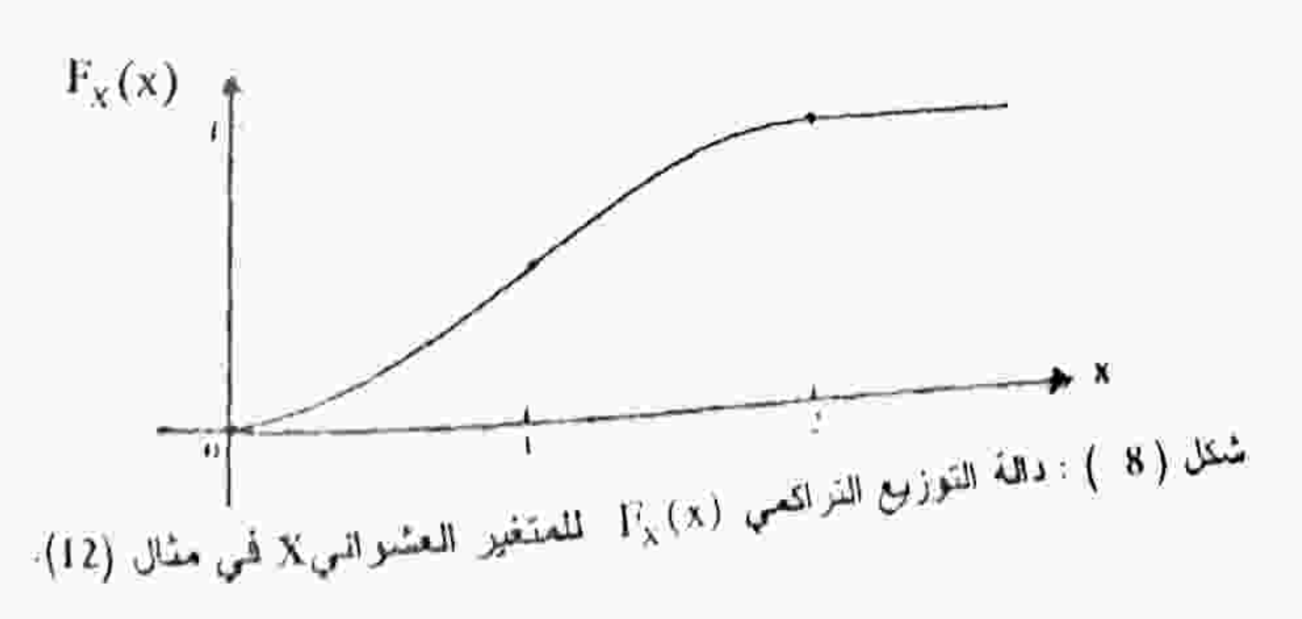
$$F_X(x) = \int_{-\infty}^{x} f_X(y)\,dy = \int_{-\infty}^{0} 0\,dy + \int_{0}^{1} y\,dy + \int_{1}^{2} (2-y)\,dy + \int_{2}^{x} 0\,dy$$

$$= 0 + \left.\frac{y^2}{2}\right|_0^1 + \left.(2y - \frac{y^2}{2})\right|_1^2 + 0 = 1$$

وبالتالي فإن دالة التوزيع التراكمي هي :

$$F_X(x) = \begin{cases} 0 & , x \leq 0 \\ \frac{x^2}{2} & , 0 < x \leq 1 \\ 2x - \frac{x^2}{2} - 1 & , 1 < x \leq 2 \\ 1 & , x \geq 2 \end{cases}$$

والشكل الذي يمثلها مبين أدناه .



شكل (8) : دالة التوزيع التراكمي $F_X(x)$ للمتغير العشوائي X في مثال (12).

جـ - الاحتمال $P(0.3 < X \le 1.5)$ يتم حسابه كالآتي :

$$P(0.3 < X \le 1.5) = P(X \le 1.5) - P(X \le 0.3)$$
$$= F_X(1.5) - F_X(0.3)$$
$$= 0.875 - 0.045 = 0.830$$

مثال (13) : إذا كان X متغيراً عشوائياً بدالة توزيع تراكمي (c. d. f.) معرفة كما يلي :

$$F_X(x) = \begin{cases} 0 & , x < 0 \\ \dfrac{x^3}{8} & , 0 \le x \le 2 \\ 1 & , x > 2 \end{cases}$$

أوجد دالة كثافة الاحتمال (p.d.f.) .

الحـل :

حيث أن $F_X(x)$ معرفة بشكل مختلف في حالات مختلفة ، وعليه يجب دراسة كل حالة على حده وبالتالي فإن

$$\frac{dF_X(x)}{dx} = \frac{d(0)}{dx} = 0 \qquad , x < 0$$

$$\frac{dF_X(x)}{dx} = \frac{d}{dx}\left(\frac{x^3}{8}\right) = \frac{3x^2}{8} \qquad , 0 \le x \le 2$$

$$\frac{dF_X(x)}{dx} = \frac{d(1)}{dx} = 0 \qquad , x > 0$$

وعليه فإن دالة كثافة الاحتمال ستكون كما يلي :

$$f_X(x) = \begin{cases} \dfrac{3x^2}{8} & , 0 \le x \le 2 \\ 0 & , ow. \end{cases}$$

وأخيراً ننهي هذا النقاش بملحوظتين ، الأولى أنـه فـي معظـم الحـالات مـن الممكـن أن يكـون المتغير العشوائي X متصلاً أو منفصلاً ، ولكن من الممكن فـي الواقـع العملـي وجـود خلـيط مـن النوعين ، وبالتالي إن أي دالة توزيع تراكمي يمكن تحليلها إلى جزئين وذلك على النحو الآتي :

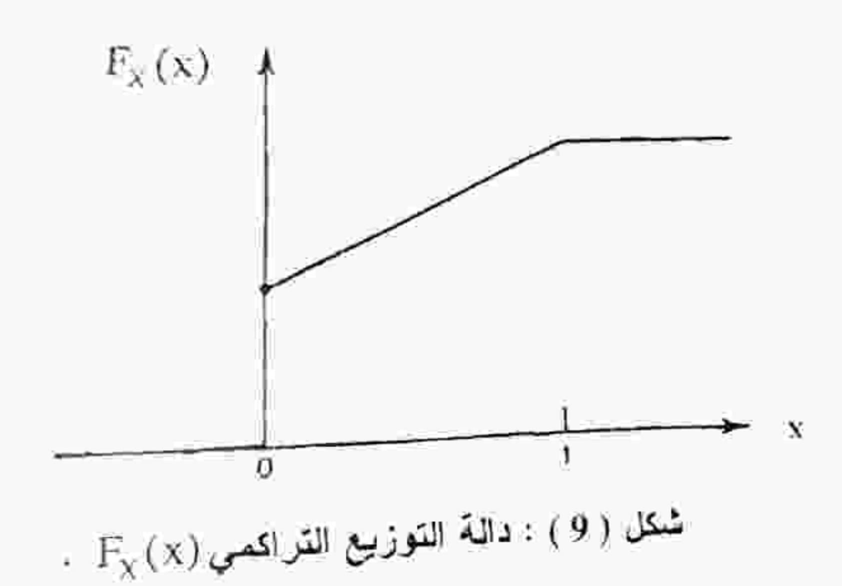
$$F_X(x) = a\,F_d(x) + (1-a)\,F_c(x) \qquad , 0 \le a \le 1$$

حيث $F_c(.)$ و $F_d(.)$ دالتا توزيع .وإن $F_c(.)$ تمثل دالة توزيع متغير عشوائي متصل ، و $F_d(.)$ تمثل دالة توزيع متغير عشوائي منفصل . وفي الواقع فإنه يمكن تحليل هذه الدالة إلى أكثر من ذلك ولكن سوف لن نتعرض لهذا التحليل هنا .

مثال (14): بفرض أن X متغيراً عشوائياً بدالة توزيع تراكمي

$$F_X(x) = \begin{cases} 0 & , x < 0 \\ \frac{1}{2} & , x = 0 \\ \frac{(x+1)}{2} & , 0 < x < 1 \\ 1 & , x \geq 1 \end{cases}$$

الحظ أن دالة التوزيع $F_X(x)$ لها قفزة عند $x = 0$ وإن $F_X(x)$ متصلة في الفترة $(0,1)$ كما في شكل (9) أدناه .



شكل (9) : دالة التوزيع التراكمي $F_X(x)$.

إن دالة التوزيع التراكمي للمتغير العشوائي X ليست متصلة ولا هي منفصلة ولكنها خليط بينهما وعليه يمكننا كتابتها على الصورة الآتية :

$$F_X(x) = \frac{1}{2} F_d(x) + \frac{1}{2} F_c(x)$$

حيث

$$F_d(x)=\begin{cases}0 & ,x<0\\ 1 & ,x\ge 0\end{cases}$$

وإن

$$F_c(x)=\begin{cases}0 & ,x\le 0\\ x+1 & ,0<x<1\\ 1 & ,x\ge 1\end{cases}$$

إذن $F_c(x)$ دالة توزيع تراكمي بدالة كثافة احتمال

$$f_c(x)=\begin{cases}1 & ,0<x<1\\ 0 & ,ow.\end{cases}$$

أما الملحوظة الثانية فهي تتعلق بالتعريف الآتي .

تعريف (7): نقول أن المتغير العشوائي X متماثل حول النقطة a إذا تحقق الشرط التالي :

$P(X\ge a+x)=P(X\le a-x)$ لجميع قيم x .

إضافة إلى ذلك فإننا نقول إن دالة التوزيع التراكمي $F_X(x)$ للمتغير العشوائي X متماثلة حول النقطة a إذا تحقق الشرط التالي :

$$F_X(a-x)=1-F_X(a+x)+P(X=a+x)\ .$$

وعلى وجه الخصوص إذا كان X متغيراً عشوائيا متصلاً دالة كثافة احتماله $f_X(.)$ فهو متماثل حول a إذا وفقط إذا كانت $f_X(a-x)=f_X(a+x)$ لجميع قيم x . وفي حالة $a=0$ فإننا نقول بأن $f_X(.)$ أو $F_X(.)$ أو X متماثلة .

ومن الأمثلة على ذلك ، دالة كتلة الاحتمال $P(X=1)=\frac{1}{2}=P(X=0)$ متماثلة حول $\frac{1}{2}$ ،

ودالة كثافة الاحتمال $f_X(x)=\frac{1}{\sqrt{2\pi\sigma^2}}e^{\frac{-(x-\mu)^2}{2\sigma^2}}\quad ,x\in R$ متماثلة حول μ .

5-3 التوقع الرياضي Mathematical Expectation

1-5-3 القيمة المتوقعة The expected value

من المفاهيم المهمة والمفيدة في المسائل التي تتعلق بالمتغيرات العشوائية هو التوقع الرياضي، وفي هذا البند سنعرض تعريفات ونتائج تتعلق بهذا الموضوع. وسنفترض أن X متغير عشوائي ، وإن $g(.)$ دالة ذات قيمة حقيقية معرفة على Ω حيث : $g(X)=g(X(\omega))$ و

$\omega \in \Omega$ وبذلك ستكون $g(X)$ هي الأخرى متغيرا عشوائيا في جميع الحالات التي سنتعرض إليها في هذا البند .

تعريف (8) : إذا كان X متغيرا عشوائيا بدالة كتلة احتمال $p_X(.)$ فإننا نعرف متوسط(X)g أو القيمة المتوقعة (The expected value) لها بالعلاقة الآتية :

$$E(g(X)) = \sum_j g(x_j)\, p_X(x_j)$$

بشرط أن $\sum_{\forall j} |g(x_j)|\, p_X(x_j) < \infty$.

أما إذا كان X متغيرا عشوائيا متصلاً بدالة كثافة احتمال $f_X(.)$ فإن :

$$E(g(X)) = \int_{-\infty}^{\infty} g(x)\, f_X(x)\, dx$$

بشرط أن $\int_{-\infty}^{\infty} |g(x)|\, f_X(x)\, dx < \infty$.

إذن على ضوء هذا التعريف فإن (g(X تكون موجودة إذا تحقق الشرط $E|g(X)| < \infty$.

ملحوظة : إذا كانت $g(X) = X$ فإن $E(g(X)) = E(X)$ ويطلق على (E(X القيمة المتوقعة للمتغير العشوائي X ،وهي عبارة عن **مقياس موقع** (Measure of Location) مقاس بنفس وحدات قياس المتغير العشوائي X ،ويعبر عن قيم المتغير المعرفة على توزيع احتمال X بقيمة واحدة تعطي بعض المعلومات عن موقع التوزيع الاحتمالي . وغالبا ما يرمز للقيمة المتوقعة بالرمز μ ،أي أن $E(X) = \mu$ وطريقة حسابها تعتمد على نوعية المتغير العشوائي وذلك كالآتي

أ - إذا كان X متغيرا عشوائيا منفصلاً فإن

$$E(X) = \sum_x x\, p_X(x) \qquad (6)$$

ب- إذا كان X متغيرا عشوائيا متصلاً فإن

$$E(X) = \int_{-\infty}^{\infty} x\, f_X(x)\, dx \qquad (7)$$

مثال (15) : إذا كان X متغيراً عشوائياً بدالة كثافة احتمال معرفة كما يلي :

$$f_X(x) = \begin{cases} 2x & ,0 < x < 1 \\ 0 & ,ow. \end{cases}$$

فأوجد $E(X)$.

الحل :

من التعريف نجد أن

$$E(X) = \int_0^1 x f_X(x)\,dx = \int_0^1 x(2x)\,dx$$

$$= \int_0^1 2x^2\,dx = 0.667$$

مثال (16) : إذا كان X متغيراً عشوائياً بدالة كتلة احتمال معرفة كما يلي :

$$p_X(x) = \begin{cases} \frac{x}{6} & ,x = 1,2,3 \\ 0 & ,ow. \end{cases}$$

أوجد $E(X)$.

الحل :

من التعريف نجد أن

$$E(X) = \sum_{x=1}^{3} x\,p_X(x) = \sum_{x=1}^{3} x\left(\frac{x}{6}\right)$$

$$= \sum_{x=1}^{3} \frac{x^2}{6}$$

$$= \frac{1}{6} + \frac{4}{6} + \frac{9}{6} = 2.333$$

مثال (17) : بفرض أن X متغير عشوائي بدالة كثافة احتمال كالآتي :

$$f_X(x) = \begin{cases} \frac{1}{x^2} & ,1 \le x < \infty \\ 0 & ,ow. \end{cases}$$

أوجد $E(X)$ إن وجدت .

الحـل :

من التعريف أعلاه نجد أن

$$E(X)=\int_1^{\infty} x f_X(x)\,dx=\int_1^{\infty} x\frac{dx}{x^2}=\int_1^{\infty}\frac{dx}{x}$$
$$=\infty$$

أي أن $E(X)$ غير موجودة وأيضًا يمكننا القول بـأن متوسـط المتغـير العشـوائي X غـير محـدود

،حيث أن التكامل الذي يعرف المتوسط هنا غير محدود .

نظرية (2) : إذا كان X متغيرًا عشوائيًا وكانت $E(X)$ موجودة فإن :

1- $E(aX+b)=aE(X)+b$ حيث أن a و b عددان حقيقيان .

2- $E(X-E(X))=0$

3- إذا وجد مقدار ثابت a بحيث $P(X\geq a)=1$ فإن $E(X)\geq a$.وإذا وجد مقدار ثابت b بحيث $P(X\leq b)=1$ فإن $E(X)\leq b$.

البرهان :

سوف نفترض أن X متغير عشوائي متصل، بدالة كثافة احتمـال $f_X(.)$ حيـث سيكون البرهان بنفس الكيفية عندما يكون X متغيرًا عشوائيًا منفصلاً، وذلك عن طريـق اسـتبدال التكامل بالمجموع . إذن

-1

$$E(aX+b)=\int_{-\infty}^{\infty}(ax+b)f_X(x)dx$$
$$=a\int_{-\infty}^{\infty}x f_X(x)dx+b\int_{-\infty}^{\infty}f_X(x)dx$$
$$=aE(X)+b$$

-2 بوضع $a=1$ و $b=-E(X)$ في (1) نحصل على النتيجة .

-3 لنفرض أولاً أن $P(X\geq a)=1$ وعليه فإن

$$E(X)=\int_{-\infty}^{\infty} x\, f_X(x)\,dx=\int_{a}^{\infty} x\, f_X(x)dx$$

$$\geq \int_{a}^{\infty} a\, f_X(x)\,dx = a\,P(X\geq a)=a.1=a$$

أما برهان الجزء الآخر فهو بالمثل .

إذن من (3) نلحظ أنه إذا كـان $P(a\leq X\leq b)=1$ فـإن $a\leq X\leq b$. ويمكن إثبات أنه إذا كان $P(X\geq a)=1$ و $E(X)=a$ فإنه يجب أن يكون صحيحًا بأن $P(X>a)=0$ وأن $P(X=a)=1$.

في بعض الأحيان يحدث سوء فهم لتعبير (القيمة المتوقعة) ، وذلك لأن $E(X)$ ليست بالضرورة قيمة المتغير العشوائي X المتوقعة عند أجراء التجربة . ومـن الأمثلة على ذلك عند إلقاء زهرة نرد مرة واحدة فإن القيمة المتوقعة هي 3.5 وهي ليست إحـدى القيم الممكنة للمتغير العشوائي X. لقد كان أحد تفسيراتنا للاحتمال مبنـي على مفهوم التكرار النسبي وذلك كنهاية للتكرارات النسبية المشاهدة في المحاولات المتكررة ، وبناء ًا على هذا المفهوم يمكن التفكير فـي القيمة المتوقعة على أنها نهاية للمتوسطات $\left(\frac{X_1+X_2+...+X_n}{n}\right)$ كلما اقتربت n من ∞ ، حيث $X_1, X_2, ..., X_n$ تمثل n من المشاهدات للمتغير العشوائي X عند تكرار التجربة n من المـرات . وهذا يعني أنه إذا تم إلقاء زهرة النرد n من المـرات مثلا ، فـإن متوسط قيم النقاط المشـاهدة على الزهرة ، عندما تكون n كبيرة من المرجح أن يكون قريبًا من 3.5 . وبعبـارة أخرى عندما تكون n كبيرة فإن المتوسط $\left(\frac{X_1+X_2+...+X_n}{n}\right)$ يكون قريبًا من $E(X)$.وأخيرًا نـود أن نشير هنا إلى أن قيمة $\mu=E(X)$ يجب أن تكون داخل نطاق الدالة $f_X(.)$ أو $p_X(.)$. وعندما يكون المتغير العشوائي منفصلا فإن قيمة μ يجب أن يكون أكبر من أصغر قيمـة واقل مـن أكبر قيمة ممكنة للمتغير العشوائي .

مثال (18) : بفرض أن

$$f_X(x)=\begin{cases}\frac{1}{5} & ,0<x<5\\ 0 & ,ow.\end{cases}$$

تمثل دالة كثافة الاحتمال للمتغير العشوائي X .أوجد

أ- $E(X)$ و $E(3X+2)$

ب- $P(X > E(X))$

الحـل :

أ- من تعريف التوقع الرياضي نجد أن

$$E(X) = \int_0^5 x\, f_X(x)\, dx = \int_0^5 x\left(\frac{dx}{5}\right)$$

$$= \frac{1}{5}\int_0^5 x\, dx = \frac{5}{2} = 2.5$$

ومن (1) بالنظرية السابقة نجد أن

$$E(3X+2) = 3E(X) + 2 = 3(2.5) + 2 = 9.5$$

ب-

$$P(X > E(X)) = P(X > 2.5) = \int_{2.5}^5 f_X(x)\, dx$$

$$= \int_{2.5}^5 \frac{1}{5} dx = 0.5$$

أخيرًا سوف ننهى هذا البند بملحوظتين هما :

ا- إذا كـان X متغيراً عشـوائياً بتوزيـع متمـاثل حـول النقطـة a ،وكـانت $E(X)$ موجـودة فـإن $E(X)= a$.

إن برهان هذه الخاصية متأتيا من الحقيقة وهى أن X-a و a-X لهما نفس التوزيع ،وعليه فإن $E(X-a)= E(a-X)$ ومنها $E(X) - a = a - E(X)$ وبالتالي فإن $E(X) = a$.

ب- إذا كانت القيم الممكنة للمتغير العشوائي X هي 2,1,0,... فإنه توجد طريقة مفيدة في حساب القيمة المتوقعة للمتغير العشوائي X بدلالة الاحتمالات وهى :

(8)

$$E(X) = \sum_{i=0}^{\infty} P(X > i)$$

$$= P(X > 0) + P(X > 1) + P(X > 2) + \cdots$$

الحظ أن الطرف الأيمن للمعادلة أعلاه يمكن إعادة كتابته كالآتي :

$$
\begin{aligned}
P(X=1) + P(X=2) + P(X=3) + \cdots \\
+ P(X=2) + P(X=3) + \cdots \\
+ P(X=3) + \cdots \\
\vdots
\end{aligned}
$$

وبجمع الأعمدة نحصل على الآتي :

$$P(X=1) + 2P(X=2) + 3P(X=3) + \ldots = \sum_{i=0}^{\infty} i\,P(X=i) = E(X)$$

مثال (19) : إذا كان X متغيراً عشوائياً بدالة كتلة احتمال معرفة كالآتي :

$$p_X(x) = P(X=x) = \begin{cases} \dfrac{1}{2^x} & , x = 1,2,3,\ldots \\ 0 & , ow. \end{cases}$$

فأوجد $E(X)$.

الحـل :

حيث أن

$$
\begin{aligned}
P(X > x) &= \left(\frac{1}{2}\right)^{x+1} + \left(\frac{1}{2}\right)^{x+2} + \ldots \\
&= \left(\frac{1}{2}\right)^{x+1} (1 + 0.5 + (0.5)^2 + \cdots) \\
&= \left(\frac{1}{2}\right)^{x} \qquad , x = 1,2,3,\cdots
\end{aligned}
$$

وعليه من الملحوظة (ب) نجد أن

$$E(X) = 1 + 0.5 + (0.5)^2 + \cdots = 2$$

3 - 5 - 2 التبـــاين Variance

إن التباين هو مقياس لدرجة تشتت التوزيع الاحتمالي لمتغير عشوائي، وعليه إذا كانت قيمة التباين صغيرة فهي مؤشر على أن التوزيع الاحتمالي متمركز حول μ، أما إذا كانت قيمته كبيرة فهي مؤشر على أن التوزيع الاحتمالي متشتت حول μ ، ولكن من الممكن أن نجعل هذا التباين كبيرا وذلك من خلال وضع قيم احتمالية بعيدة جدا عن نقطة الأصل حتى ولو كانت هذه القيم الاحتمالية صغيرة .

تعريف (٩) : إذا كان X متغيرًا عشوائيًا وكانت $\mu = E(X)$ فإننا نعرف تباين المتغير العشوائي X (أو توزيع X) ،والذى يرمز له بالرمز $V(X)$ أو σ^2 بالعلاقة الآتية :-

$$V(X) = E(X-\mu)^2$$

وطريقة حسابه تعتمد على نوعية المتغير العشوائى وذلك كالآتى :

أ- إذا كان X متغيرًا عشوائيًا منفصلاً فإن

$$V(X) = E(X-\mu)^2 = \sum_{\forall j} (x_j - \mu)^2 p_X(x_j) \qquad (٩)$$

ب- إذا كان X متغيرًا عشوائيًا متصلاً فإن

$$V(X) = E(X-\mu)^2 = \int_{-\infty}^{\infty} (x-\mu)^2 f_X(x)\,dx \qquad (10)$$

يتضح من التعريف أعلاه أن $V(X)$ قيمة غير سالبة أي أن $V(X) \geq 0$ ، وإن الجذر التربيعي للتباين يسمى بالانحراف المعياري (standard deviation) ويرمز له بالرمز σ ، أي أن $\sigma = \sqrt{V(X)}$ ، وهو أيضًا مقياس لتشتت قيم المتغير العشوائي ويفضل استخدامه في كثير من التطبيقات عن التباين حيث أنه يستخدم نفس وحدات القياس المستخدمة للمتغير العشوائي X

خواص التباين Properties of the variance :

إذا كان X متغيرًا عشوائيًا وكانت $E|X|^2 < \infty$ فإن

1 - $V(X) = 0$ إذا وفقط إذا وجد مقدار ثابت c بحيث $P(X = c) = 1$.

البرهان :

أولاً سنفترض وجود مقدار ثابت c بحيث $P(X = c) = 1$ إذن $E(X) = c$ وأن $P[(X-c)^2 = 0] = 1$ وعليه فإن $V(X) = E(X-c)^2 = (c-c)^2 = 0$.

والآن العكس ، سوف نفترض بأن $V(X) = 0$ إذن $P(X = \mu) = 1$ وفى الواقع ، إذا افترضنا أن X متغير عشوائي منفصل وكان $P(X = x) > 0$ و $x \neq \mu$ فإن

$$V(X) = \sum_{i} (x_i - \mu)^2 P(X = x_i)$$

$$= (x-\mu)^2 P(X = x) + \sum_{i,\, x_i \neq x} (x_i - \mu) P(X = x_i) > 0$$

وهذا التناقض يبين أنه لا يمكن أن تكون $x \neq \mu$ بحيث $P(X = x) > 0$ ، وعليه فإن $P(X = \mu) = 1$ ، ويمكن البرهان بنفس الطريقة في حالة المتغير العشوائي المتصل .

2- إذا كان a و b عددان حقيقيان وكانت $g(x) = a x + b$ فإن $V(g(X)) = a^2 V(X)$.

البرهان :

إذا كانت $E(X) = \mu$ فإن $E(g(X)) = E(aX + b) = a\mu + b$ وعليه فإن

$$V(g(X)) = V(aX + b) = E\{[(aX + b) - E(aX + b)]^2\}$$
$$= E\{(aX + b - a\mu - b)^2\}$$
$$= E\{[a(X - \mu)]^2\}$$
$$= a^2 E\{(X - \mu)^2\}$$
$$= a^2 V(X)$$

إذن من هذه الخاصية يتضح أنه ، إذا كانت $a = 1$ فإن $V(X + b) = V(X)$ لأي ثابت b ، وهذه النتيجة توضح أن تغير موقع توزيع المتغير العشوائي X مسافة b وحدة على امتداد الخط الحقيقي سوف يتغير متوسط التوزيع بمقدار b وحدة ولكن لا يتأثر تشتت التوزيع حول متوسطه. أيضًا من هذه الخاصية نرى أنه إذا كانت $a = -1$ و $b = 0$ فإن $V(-X) = V(X)$ وهى تعكس التوزيع بالكامل بالنسبة إلى نقطة الأصل من الخط الحقيقي حيث ستؤدى إلى توزيع جديد وهو عبارة عن المرآة الخيالية للنقطة الأصلية . وكما رأينا سابقا فإن المتوسط سوف يتغير من μ إلى $-\mu$ ولكن أجمالي تشتت التوزيع حول متوسطه لن يتأثر .

3 - لأي متغير عشوائي X ، يكون $V(X) = E(X^2) - (E(X))^2$.

البرهان :

ألحظ أولا أنه إذا كانت $E(X^2)$ موجودة فإن $E(X)$ موجودة أيضًا ، وبوضع $\mu = E(X)$ ومن تعريف التباين نجد أن :

$$V(X) = E[(X - \mu)^2]$$
$$= E[X^2 - 2\mu X + \mu^2]$$
$$= E(X^2) - 2\mu E(X) + \mu^2$$
$$= E(X^2) - \mu^2$$

وهى صيغة أسهل في الاستخدام عند حساب التباين .

4- إذا كانت $g(x) = \frac{1}{\sigma}(x - \mu)$ فإن $E(g(X)) = 0$ و $V(g(X)) = 1$.

البرهان :

$$E(g(X)) = E\left(\frac{X-\mu}{\sigma}\right) = \frac{1}{\sigma}E(X-\mu)$$

$$= \frac{1}{\sigma}(\mu-\mu) = 0$$

ومن الخاصية (2) نجد أن

$$V(g(X)) = V\left(\frac{X-\mu}{\sigma}\right) = \frac{1}{\sigma^2}V(X-\mu) = \frac{\sigma^2}{\sigma^2} = 1$$

إن المتغير العشوائي $g(X) = \frac{1}{\sigma}(X-\mu)$ يسمى بمعاير X (standardized X) وعادةً ما يرمز للمتغير العشوائى المعاير بالرمز Z اى أن $Z = \frac{X-\mu}{\sigma}$. وهى عملية مفيدة لمقارنة توزيعين أو أكثر .

مثال (20) :إذا علمت أن متوسط درجات امتحان في مادة الرياضة يساوى 65 وبانحراف معياري يساوى 10 .أوجد

أ- الدرجة المعيارية المناظرة للدرجتين 60 و80 .

ب- الدرجة المناظرة للدرجتين المعياريتين -1 و2 .

الحـل :

أ- حيث أن $\mu = 65$ و $\sigma = 10$ وعليه فإن الدرجة المعيارية المناظرة للدرجتين 60 و80 تكون كما يلي :

$$z = \frac{80-65}{10} = 1.5 \quad و \quad z = \frac{60-65}{10} = -0.5$$

ب- لإيجاد الدرجة المناظرة لدرجة معيارية معينة يجب حل المعادلة $Z = \frac{X-\mu}{\sigma}$ بالنسبة إلى X وذلك كما يلي :

$$Z = \frac{X-\mu}{\sigma} \Rightarrow Z\sigma = X-\mu \Rightarrow X = \mu + Z\sigma$$

وعليه فإن الدرجة المناظرة للدرجتين المعياريتين -1 و 2 تكونا على الترتيب كما يلي :

$$x = 65 + (-1)(10) = 55$$

و

$$x = 65 + (2)(10) = 85$$

مثال (21) : إذا اكان X متغيرا عشوائيا بدالة كثافة احتمال معرفة كالآتي :-

$$f_X(x) = \begin{cases} 1-|x| & ,|x|<1 \\ 0 & ,ow. \end{cases}$$

أوجد تباين المتغير العشوائي X وكذلك انحرافه المعياري .

الحـل :

الحظ أولاً أن دالة كثافة الاحتمال متماثلة حول $x = 0$ وعليه فإن $E(X) = 0$ وبالتالي فإن

$$\begin{aligned} V(X) &= E(X^2) - (E(X))^2 \\ &= E(X^2) \\ &= \int_{-1}^{1} x^2 (1-|x|)\, dx \\ &= 2\int_{0}^{1} x^2 (1-x)\, dx = \frac{1}{6} \end{aligned}$$

وبذلك يكون الانحراف المعياري هو $\sigma = \sqrt{\frac{1}{6}} = 0.408$

مثال (22) : إذا كان X متغيرا عشوائيا يأخذ القيم 4 , 3 , 1 , 0 , 2- باحتمالات متساوية أوجد الانحراف المعياري والتباين للدالة $g(x) = 4x - 7$.

الحـل :

حيث أن قيم المتغير العشوائي X لها احتمالات متساوية ، وعليه فإن دالة كتلة احتماله ستكون كالآتي :-

$$p_X(x) = \begin{cases} \frac{1}{5} & ,x = -2,0,1,3,4 \\ 0 & ,ow. \end{cases}$$

وبالتالي فإن

$$E(X)=\sum_x x\,p_X(x)$$

$$=(-2)\left(\frac{1}{5}\right)+(0)\left(\frac{1}{5}\right)+(1)\left(\frac{1}{5}\right)+(3)\left(\frac{1}{5}\right)+(4)\left(\frac{1}{5}\right)=1.2$$

وأن

$$E(X^2)=\sum_x x^2\,p_X(x)$$

$$=(-2)^2\left(\frac{1}{5}\right)+(0)^2\left(\frac{1}{5}\right)+(1)^2\left(\frac{1}{5}\right)+(3)^2\left(\frac{1}{5}\right)+(4)^2\left(\frac{1}{5}\right)=6$$

ومنهما نجد أن

$$V(X)=E(X^2)-(E(X))^2$$

$$=6-(1.2)^2=4.56$$

وحيث أن $g(x)=4x-7$ وعليه من الخاصية (2) نجد أن :

$$V(g(X))=V(4X-7)=16\,V(X)=16\,(4.56)=72.96$$

وإن الانحراف المعياري للدالة $g(x)$ هو $\sigma_g=\sqrt{72.97}=8.54$.

مثال (23) : إذا افترضنا أن X متغير عشوائي بدالة كتلة احتمال معرفة كما يلي :-

$$p_X(x)=\begin{cases}1-\frac{1}{b} & ,x=0\\ \frac{1}{b} & ,x=b\\ 0 & ,ow.\end{cases}$$

حيث $b>1$. أوجد $V(X)$ عندما b=10 و b=100 . ماذا تستنتج ؟

الحل :

حيث أن

$$E(X)=(0)(1-\frac{1}{b})+(b)(\frac{1}{b})=1$$

وأن

$$V(X)=\sum_x (x-1)^2\,p_X(x)=(0-1)^2(1-\frac{1}{b})+(b-1)^2(\frac{1}{b})=b-1$$

وعليه عندما $b = 10$ نجد أن

$$P_X(x) = \begin{cases} 0.9 & , x = 0 \\ 0.1 & , x = 10 \end{cases}$$

وبالتالي فإن $E(X) = 1$ و $V(X) = 9$

ولكن عندما $b = 100$ فإن

$$P_X(x) = \begin{cases} 0.99 & , x = 0 \\ 0.01 & , x = 100 \end{cases}$$

وبالتالي فإن $E(X) = 1$ و $V(X) = 99$.

إذن يتضح أنه كلما زادت قيمة b ، فإن معظم الاحتمال يتمركز عند الصفر أى عند القيمة $X = 0$ ،ويقل هذا الاحتمال عند $X = b$ وهكذا فإن X ستكون أقل تغيرًا . ولكن تباين X يزداد كلما زادت قيمة b وهذا يعنى أن X أكثرا تغيرا . والسبب في هذا التناقض يرجع إلى أن، التباين لا يكون حساسا للاحتمالات الصغيرة التي تكون بعيدة عن الجزء الرئيسي في التوزيع الاحتمالي إذن يتضح من هذا المثال أن التباين ليس مؤشرا جيدًا على شكل التوزيع في جميع الأحوال.

3 - 5 - 3 متباينة تشيبشيف Chebyshev's Inequality

إذا كان X متغيرًا عشوائيًا وكانت $E(X)$ موجودة وكذلك $V(X)$ موجود فإنهما لا يعطيان معلومات كافية عن هذا المتغير العشوائي ،ولكن يمكن وضع تخمين عن بعض خواصه والذي من الممكن أن يكون مفيد ، هذا التخمين يتم من خلال استخدام متباينة تشيبشيف .

نظرية (3) : إذا كان X متغيرًا عشوائيًا بحيث $P(X \geq 0) = 1$ (أي أنه يأخذ قيم غير سالبة) فإنه لأي عدد k حيث $k > 0$ يكون

$$P(X \geq k) \leq \frac{E(X)}{k} \qquad (11)$$

البرهان :

سوف نفترض بأن X متغير عشوائي متصل بدالة كثافة احتمال $f_X(.)$ ،وعليه فإن

$$E(X) = \int_{-\infty}^{\infty} x f_X(x)\,dx$$

$$= \int_{0}^{k} x f_X(x)\,dx + \int_{k}^{\infty} x f_X(x)\,dx$$

$$\geq \int_{k}^{\infty} x f_X(x)\,dx$$

$$\geq \int_{k}^{\infty} k f_X(x)\,dx = k\int_{k}^{\infty} f_X(x)\,dx$$

$$= k\,P(X \geq k)$$

بالقسمة على k ،ومنها النتيجة .وسيكون البرهان بالمثل في حالة ما يكون X متغيرًا عشوائيًا متصلاً .

نظرية (4) : متباينة تشيبشيف

إذا كان X متغيرًا عشوائيًا وكان $V(X) = \sigma^2$ حيث $0 < \sigma < \infty$ فإنه لأى عدد $k > 0$ يكون

$$P(|X-\mu| \geq k\sigma) \leq \frac{V(X)}{k^2\sigma^2} = \frac{1}{k^2} \qquad (12)$$

أو

$$P(|X-\mu| \geq k) \leq \frac{V(X)}{k^2} = \frac{\sigma^2}{k^2}$$

البرهان :

حيث أن $(X-\mu)^2$ متغير عشوائى غير سالب ، وعليه بتطبيق النظرية السابقة نجد أن

$$P[(X-\mu)^2 \geq k^2\sigma^2] \leq \frac{E[(X-\mu)^2]}{k^2\sigma^2}$$

وحيث أن $(X-\mu)^2 \geq k^2\sigma^2$ إذا وفقط إذا كان $|X-\mu| \geq k\sigma$ ،وعليه فإن المعادلة اعلاه مكافئة إلى

$$P[|X-\mu| \geq k\sigma] \leq \frac{E[(X-\mu)^2]}{k^2\sigma^2} = \frac{\sigma^2}{k^2\sigma^2} = \frac{1}{k^2}$$

وبهذا يكتمل البرهان .

إن النظريتين السابقتين تساعدان في إيجاد أو وضع حدود على الاحتمالات ،عندما يكون المتوسط أو المتوسط والتباين للتوزيع الاحتمالي معلومين .بالطبع إذا كان التوزيع الاحتمالي معلومًا ، فإن الاحتمالات المرغوب فيها يمكن حسابها مباشرة ولسنا بحاجة لوضع حدود عليها .

ملحوظة : إذا كان X متغيرًا عشوائيًا وبتباين محدود فإن

$$P(|X-\mu| < k\sigma) \geq 1-\frac{1}{k^2} \qquad (13)$$

إن الملحوظة عبارة عن إعادة كتابة للمعادلة الأولى بالنظرية .ونلحظ منها أنه كلما كانت k كبيرة كان احتمال انحراف X عن μ بمقدار أكبر من أو يساوى $k\sigma$ ، صغيرًا . إن النظرية والملحوظة ينطبقان على جميع التوزيعات التي يكون فيها $E(X^2) < \infty$ وسوف نستخدم متباينة تشيبشيف في برهان قانون الأعداد الكبيرة . لاحظ أيضًا أن هذه المتباينة تتضمن ثلاث معلمات وهى μ و σ و k وبمعلومية هذه المعالم فإن المتباينة تقول بأن X تنتمي للفترة $(\mu-k\sigma \ ،\mu+k\sigma)$ باحتمال قدره على الأقل $1-\frac{1}{k^2}$. ومن الناحية العملية فإن μ و σ غالبًا ما يكونان مجهولين، وعليه وكما أشرنا أعلاه نحاول أيجاد أو وضع حدود تغطى المعلمة المجهولة μ أو σ باحتمال محدد مسبقا وليكن $1-\alpha$ مثلاً ، حيث $0 < \alpha < 1$.فإذا فرضنا أن μ غير معلومة ولكن σ معلومة ونود أيجاد قيمة كل من $a(X)$ و $b(X)$ بحيث $P(a(X) < \mu < b(X)) \geq 1-\alpha$ فسوف يكون اختيارنا هو أن $1-\alpha = 1-\frac{1}{k^2}$ وعليه فإن $k = \frac{1}{\sqrt{\alpha}} > 1$ ، وبالتالي من الملحوظة نجد أن

$$P(\mu-\frac{\sigma}{\sqrt{\alpha}} < X < \mu+\frac{\sigma}{\sqrt{\alpha}}) \geq 1-\alpha$$

$$\Leftrightarrow \quad P(X-\frac{\sigma}{\sqrt{\alpha}} < \mu < X+\frac{\sigma}{\sqrt{\alpha}}) \geq 1-\alpha$$

وهكذا نجد أن $a(X) = X-\frac{\sigma}{\sqrt{\alpha}}$ و $b(X) = X+\frac{\sigma}{\sqrt{\alpha}}$. وبالمثل إذا كانت μ معلومة و σ غير معلومة فإنه يمكن الإثبات أن $P(0 < \sigma \leq |X-\mu|\sqrt{\alpha}) \leq \alpha$.

مثال (24) : إذا كان X متغيراً عشوائياً يمثل عدد النقاط التي ستظهر عند إلقاء زهرة نرد متزنة مرة واحدة .فأوجد حداً أعلى للاحتمال $P(|X-\mu| \ge k\sigma)$ باستخدام متباينة تشيبشيف عندما $k = 2.5$.

ثم أوجد قيمة هذا الاحتمال مباشرة باستخدام دالة الاحتمال

الحل :

إنه من السهل التعرف على أن دالة كتلة احتمال المتغير العشوائي X ستكون كالآتي :

$$p_X(x) = \begin{cases} \dfrac{1}{6} & , x = 1,2,3,4,5,6 \\ 0 & , ow. \end{cases}$$

ومنها نجد أن $E(X) = 3.5$ و $V(X) = 2.917$ ومن النظرية (4) ولأي قيمة k حيث $k > 0$ يكون

$$P[|X-3.5| > k] < \frac{V(X)}{k^2}$$

وبوضع $k = 2.5$ نجد أن

$$P[|X-3.5| > 2.5] < \frac{2.917}{6.25} \approx 0.467$$

بينما الاحتمال الحقيقي هو

$$\begin{aligned} p &= P(|X-3.5| \ge 2.5] \\ &= P[(X-3.5) > 2.5] + P[(X-3.5) < -2.5] \\ &= P(X > 6) + P(X < 1) = 0 \end{aligned}$$

وذلك لأن X لا يمكن أن تكون أقل من 1 أو أكبر من 6 .

مثال (25) : إذا افترضنا أن X متغير عشوائى بدالة كثافة احتمال معرفة كما يلي :

$$f_X(x) = \begin{cases} 1 & , 0 < x < 1 \\ 0 & , ow. \end{cases}$$

أوجد حداً أعلى للاحتمال $P(|X-\mu| \ge k\sigma)$ باستخدام متباينة تشيبشيف ثم أوجد قيمة هذا الاحتمال مباشرة باستخدام الدالة المعطاة عندما $k = 2.0$.

الحل :

إنه من السهل الإثبات بأن $E(X) = \dfrac{1}{2}$ و $\sigma^2 = V(X) = \dfrac{1}{12}$ وبناءً على النظرية (4) فإن

$$P[|X-\tfrac{1}{2}| \ge k\sigma] \le \frac{V(X)}{k^2\sigma^2} = \frac{1}{k^2}$$

$$\Rightarrow \qquad P[|X-\tfrac{1}{2}| \ge \frac{k}{\sqrt{12}}] \le \frac{1}{k^2}$$

وعندما $k = 2$ فإن

$$P[|X-\tfrac{1}{2}| \ge \frac{2}{\sqrt{12}}] \le \frac{1}{4} = 0.25$$

ولكن قيمة الاحتمال الحقيقي (أى مباشرة باستخدام دالة الكثافة) هى

$$P[|X-\tfrac{1}{2}| \ge \frac{2}{\sqrt{12}}] = P[(X-\tfrac{1}{2}) \ge \frac{2}{\sqrt{12}}] + P[(X-\tfrac{1}{2}) \le -\frac{2}{\sqrt{12}}] = 0$$

مثال (26) : إذا كان X متغيراً عشوائياً بدالة كثافة احتمال معرفة كالآتي :

$$f_X(x) = \begin{cases} \frac{1}{10} & ,0 < x < 10 \\ 0 & ,ow. \end{cases}$$

أوجد حداً أعلى للاحتمال $P(|X-\mu| \ge k\sigma)$ باستخدام متباينة تشيبشيف ثم أوجد قيمة هذا الاحتمال مباشرة باستخدام الدالة المعطاة عندما $k = \frac{4\sqrt{3}}{5}$.

الحل :

مباشرة من دالة كثافة الاحتمال نجد أن $E(X) = 5$ و $V(X) = \frac{25}{3}$ وعليه من النظرية (4) نجد أن

$$P[|X-5| \ge 4] \le \frac{25}{3(16)} \approx 0.52$$

بينما قيمة هذا الاحتمال مباشرة هي

$$P[|X-5| > 4] = P[(X-5) > 4] + P[(X-5) < -4]$$
$$= P(X > 9) + P(X < 1) = 0.20$$

نلحظ من خلال هذه الأمثلة أنه بالرغم من أن متباينة تشيبشيف صحيحة ، إلا أن الحد الأعلى للاحتمال $\frac{1}{k^2}$ الذي تعطيه ليس قريباً من الاحتمال الحقيقي $P[|X-\mu| \ge k\sigma]$ ، وهذا يقترح إمكانية تطوير اللامتساويه ، ولكن إذا أردنا لهذه اللامتساوية أن تتحقق لجميع قيم $k > 0$ ولجميع

المتغيرات العشوائية التي تباينها محدود، فإن مثل ذلك التطوير غير ممكن بدون وضع افتراضات أخرى حول توزيع المتغير العشوائي ومع ذلك نجد أن اللأمتساوية مفيدة جدًا وغالبًا ما نلجأ لاستخدامها .

3 - 6 العزوم Moments

إن عزوم المتغير العشوائي (أو عزوم التوزيع الاحتمالي لمتغير عشوائي) هي القيم المتوقعة لدوال معينة بدلالة المتغير العشوائي الذي له دالة كتلة احتمال $p_X(.)$ أو دالة كثافة احتمال $f_X(.)$.

تعريف (10) : العزوم اللامركزية Noncentral Moments

إذا كان X متغيرًا عشوائيًا فإن العزم اللامركزي الرائي لهذا المتغير يرمز له بالرمز μ'_r ويعرف كما يلي : $\mu'_r = E(X^r) \quad , r = 1,2,3,...$. وتعتمد طريقة حسابه على نوعية المتغير العشوائي .

1- إذا كان X متغيرًا عشوائيًا منفصلاً فإن :

$$\mu'_r = E(X^r) = \sum_x x^r \; p_X(x) \qquad (14)$$

2- إذا كان X متغيرًا عشوائيًا متصلاً فإن :

$$\mu'_r = E(X^r) = \int_{-\infty}^{\infty} x^r \, f_X(x) dx \qquad (15)$$

إذا كان التوقع موجودًا ،حيث r = 1, 2, 3,.... . وعلى وجه الخصوص إذا كانت :

$$r = 1 \;\Rightarrow\; \mu'_1 = E(X) = \mu$$
$$r = 2 \;\Rightarrow\; \mu'_2 = E(X^2)$$
$$r = 3 \;\Rightarrow\; \mu'_3 = E(X^3)$$
$$\vdots \qquad\qquad \vdots$$

وهكذا .

ومن هذه العلاقات نلحظ أن $\sigma^2 = V(X) = \mu'_2 - (\mu'_1)^2$. ونقول بأن العزم r موجودًا إذا وفقط إذا كان $E(|X|^r) < \infty$. وإذا كان المتغير العشوائي X محددًا أي إذا وجدت أعداد

محدودة a و b مثلا، بحيث $P(a \le X \le b) = 1$ فإن جميع العزوم اللامركزية للمتغير العشوائي X من الضروري أن تكون موجودة .ولكن من الممكن أن تكون جميع العزوم اللامركزية متواجدة بالرغم من أن المتغير العشوائي X ليس محدود ، حيث أنه تم الإثبات كما في النظرية القادمة إذا كان العزم r للمتغير العشوائي X موجوداً فإن جميع العزوم ذات المرتبة التي أقل من r ستكون موجودة أيضا .

نظرية (5) : إذا كانت $E(|X|^r) < \infty$ حيث r عدد صحيح موجب ،فإنه لأي عدد صحيح موجب p ، حيث $p < r$ يكون $E(|X|^p) < \infty$.

البرهان :

سوف نفترض أن X متغير عشوائي متصل بدالة كثافة احتمال $f_X(.)$ ، وعليه فإن

$$E(|X|^p) = \int_{-\infty}^{\infty} |X|^p f_X(x)\,dx$$
$$= \int_{|x|\le 1} |X|^p f_X(x)\,dx + \int_{|x|>1} |X|^p f_X(x)\,dx$$
$$\le \int_{|x|\le 1} 1.f_X(x)\,dx + \int_{|x|>1} |X|^r f_X(x)\,dx$$
$$\le P(|X| \le 1) + E(|X|^r)$$

ومن معطيات النظرية أن $E(|X|^r) < \infty$ ، وعليه فإن $E(|X|^p) < \infty$.بالمثل يمكن برهان النظرية عندما يكون المتغير العشوائي X منفصلاً .

تعريف (11) : العزوم المركزية Central Moments

إذا كان X متغيرًا عشوائيًا ،فإن العزم المركزي الرائي لهذا المتغير حول النقطة a ، يرمز له بالرمز $\mu_r(a)$ ، ويعرف كما يلي :

$$\mu_r(a) = E\{(X-a)^r\} \qquad , r = 1,2,3,...$$

حيث يتم حسابه على حسب نوعية المتغير العشوائي .

1- إذا كان X متغيرًا عشوائيًا منفصلاً فإن :

$$\mu_r(a) = E\{(X-a)^r\} = \sum_x (x-a)^r p_X(x) \qquad (16)$$

2- إذا كان X متغيراً عشوائياً متصلاً فإن :

$$\mu_r(a) = E\{(X-a)^r\} = \int_{-\infty}^{\infty} (x-a)^r f_X(x)dx \qquad (17)$$

وإذا كانت $a = \mu = E(X)$ فإن العزم الرائي المركزي للمتغير العشوائي X حول μ ، والذي يرمز له بالرمز μ_r هو :

$$\mu_r = E\{(X-\mu)^r\} \qquad , r = 1,2,3,... \qquad (18)$$

وعليه إذا كانت : $r = 1$ فإن

$$\mu_1 = E\{(X-\mu)\} = 0$$

وإذا كانت : $r = 2$ فإن

$$\mu_2 = E\{(X-\mu)^2\} = E(X^2) - \mu^2$$
$$= \mu_2' - (\mu_1')^2 = \sigma^2$$

وعندما $r = 3$ فإن

$$\mu_3 = E\{(X-\mu)^3\}$$
$$= E(X^3) - 2\mu E(X^2) + 2(\mu)^3$$
$$= \mu_3' - 2\mu_1'\mu_2' + 2(\mu_1')^3$$

وهكذا لبقية قيم r . ألحظ أنه إذا كانت دالة كتلة الاحتمال أو دالة كثافة الاحتمال للمتغير العشوائي X متماثلة حول μ فإن جميع العزوم المركزية الفردية تساوى صفرا .

في بعض الأحيان ، تكون العزوم السابقة مثل μ و $\mu_2 = \sigma^2$ غير موجودة ،وفي مثل تلك الحالات فإن ما تعرضنا إليه سابقا لا يمكن تطبيقه ،وعليه سنعرض مقاييس أخرى للموقع والتشتت والتي تكون دائما موجودة .

تعريف (12) : التجزيء Quantile

التجزيء ذو المرتبة p (حيث $0 < p < 1$) للمتغير العشوائي X (أو لتوزيع المتغير العشوائي X) هو أي عدد ζ_p بحيث

(19)

أو بصورة أخرى

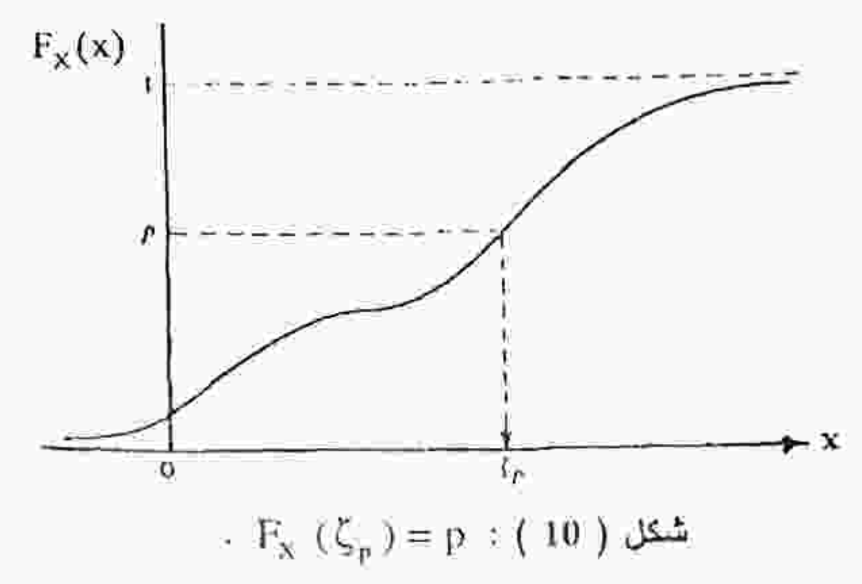
$$P(X \geq \zeta_p) \geq 1-p \quad , P(X \leq \zeta_p) \geq p$$

$$P(X < \zeta_p) \leq p \leq P(X \leq \zeta_p) \qquad (20)$$

إذا كان للمتغير العشوائي X دالة توزيع تراكمية $F_X(.)$ فإن ζ_p تمثل تجزؤا من الرتبة p إذا كانت تفي بالعلاقة الآتية :

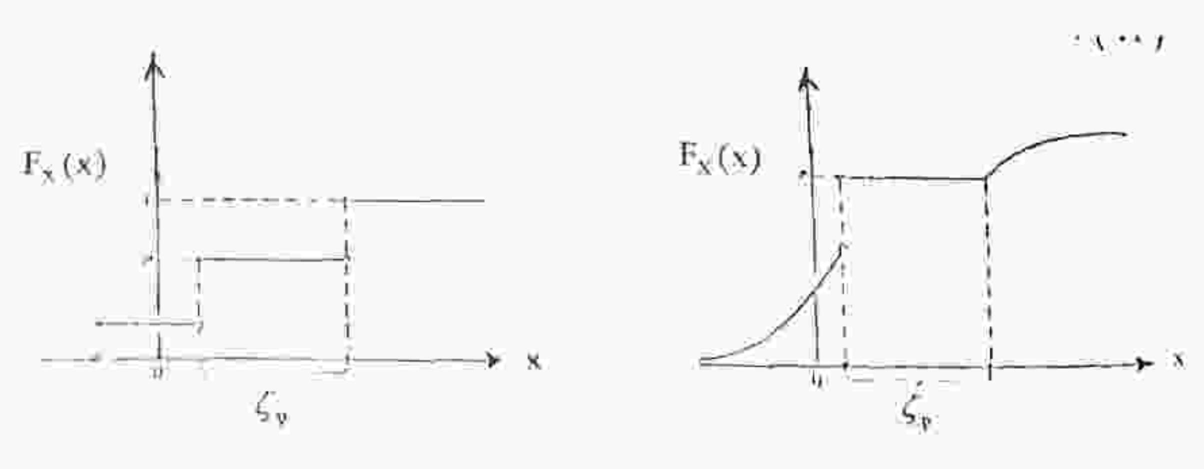
$$F_X(\zeta_p) - P(X = \zeta_p) \leq p \leq F_X(\zeta_p) \qquad (21)$$

وعلى وجه الخصوص إذا كانت $F_X(.)$ متصلة عند ζ_p فإن ζ_p تمثل حل للمعادلة $F_X(\zeta_p) = p$ كما في شكل (10) .



شكل (10) : $F_X(\zeta_p) = p$.

علاوة على ذلك قد لا يكون التجزؤ وحيدا إذا لم تكن $F_X(.)$ دالة مطردة .كما هو مبين في شكل (11) .



شكل (11) : التجزؤ قد لا يكون وحيدا .

ملحوظة :

1- إذا كانت $p = 0.5$ فإن $\zeta_{0.5}$ تمثل الوسيط (Median) ،وفي بعض المراجع يعرف الوسيط على أنه أي عدد يحقق $P(X \le \zeta_{0.5}) \ge 0.5$ و $P(X \ge \zeta_{0.5}) \ge 0.5$ وإذا كان X متغيرا عشوائيا متصلا فإن الوسيط يفي بالعلاقة الآتية :

$$\int_{-\infty}^{\zeta_{0.5}} f_X(x)\,dx = 0.5 = \int_{\zeta_{0.5}}^{\infty} f_X(x)\,dx \qquad (22)$$

2- إذا كانت $p = 0.25$ فإن $\zeta_{0.25}$ تمثل الربيع الأدنى (Lower Quartile).

3- إذا كانت $p = 0.75$ فإن $\zeta_{0.75}$ تمثل الربيع الأعلى (Upper Quartile) .

4- هناك مقياس أخر يعد من مقاييس الموقع وهو المنوال (Mode) ،ويعرف على أنه النقطة (إن وجدت) التي تكون عندها $f_X(.)$ أو $p_X(.)$ في نهايتها العظمى . وإذا كان X متغيرا عشوائيًا متصلاً فإنه يمكن الحصول على المنوال من خلال البحث عن قيمة (أو قيم X) التي تجعل الدالة $f_X(.)$ في نهايتها العظمى ،وهذا يعني البحث عن القيمة أو القيم التي تحقق المعادلة الآتية :

$$f'_X(x) = \frac{df_X(x)}{dx} = 0 \qquad (23)$$

بشرط أن $f''_X(x) < 0$.

أما إذا كان X متغيرًا عشوائيًا منفصلاً فإن المنوال هو القيمة التي تقابل أكبر احتمال .وإذا لوحظ تساوى احتمال قيمتين مختلفتين فإن التوزيع يسمى ثنائي المنوال كما في شكل (12) .



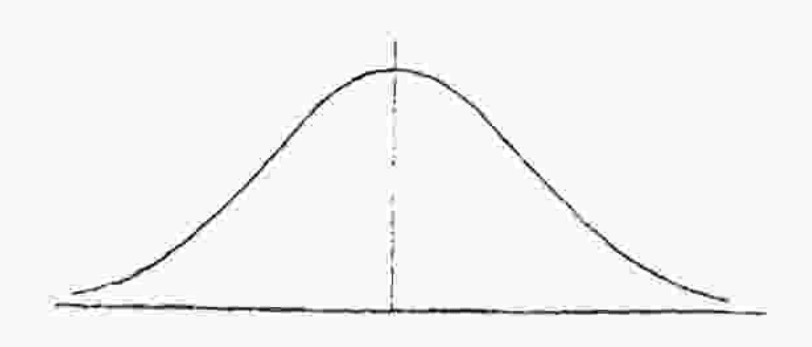
شكل (12) : توزيعات بمنوال ومنوالين .

وإذا كان لجميع القيم نفس الاحتمال فإنه يقال بأن التوزيع الاحتمالي للمتغير العشوائي X ليس له منوال .

5- الانحراف الربيعي (Quartile Deviation) : إن هذا المقياس يعد من مقاييس التشتت إلا أنه غير دقيق ولكنه مفيد في قياس التشتت في الحالات التي يتعذر فيها حساب التباين .إن وحدات قياس الانحراف الربيعي هي نفس وحدات قياس المتغير العشوائي ويرمز له بالرمز Q ، ويعرف كما يلي :

$$Q = \frac{\zeta_{0.75} - \zeta_{0.25}}{2} \qquad (24)$$

6- يستخدم العزم الثالث المركزي حول المتوسط ، μ_3 ، مقياسا لتماثل أو التواء التوزيعات .فإذا كان التوزيع متماثلاً كما في شكل (13) ، فإن $\mu_3 = 0$.



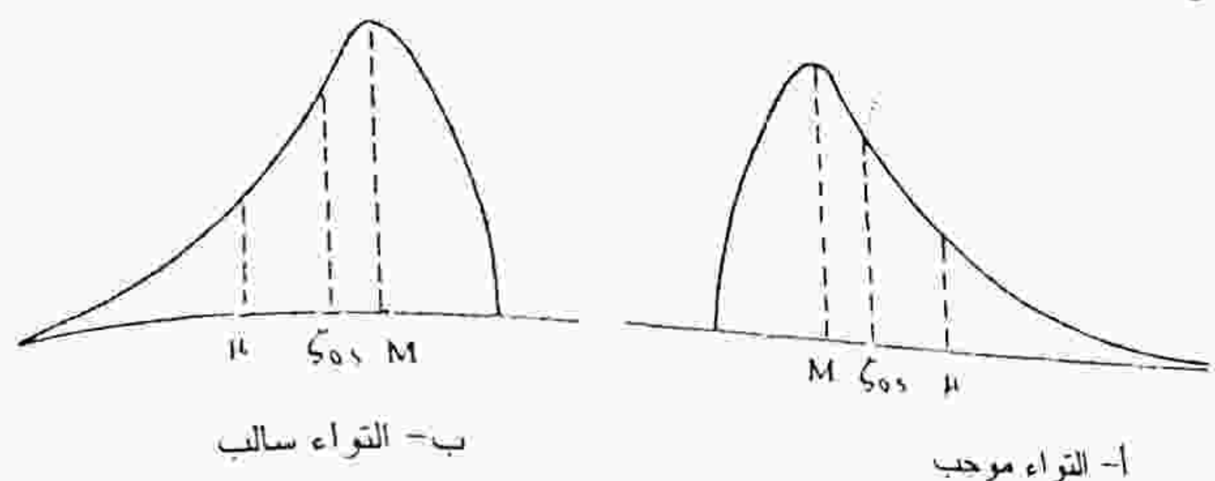
شكل (13) : توزيع متماثل

أما إذا كانت $\mu_3 \neq 0$ ،فإننا نقول بأن التوزيع ملتويا ويمكن قياس الالتواء بمقياس يسمى بمعامل الالتواء (Skewness) ، ويرمز له بالرمز γ وله الصيغة الآتية :

$$\gamma = \frac{\mu_3}{\sigma^3} \qquad (25)$$

وقد يكون الالتواء موجبًا ($\gamma > 0$) ،إذا كانت المساحة التي على يمين المنوال أكبر من تلك التي على يساره ، وفي هذه الحالة فإن $\mu > \zeta_{0.5} > M$ (حيث M تمثل المنوال) ؛ أما إذا كان

الالتواء سالب ($\gamma < 0$) فإن المساحة التي على يمين المنوال أصغر من تلك التي على يسار و
وفي هذه الحالة فإن $\mu < \zeta_{0.5} < M$ كما في شكل (14) .



ب- التواء سالب

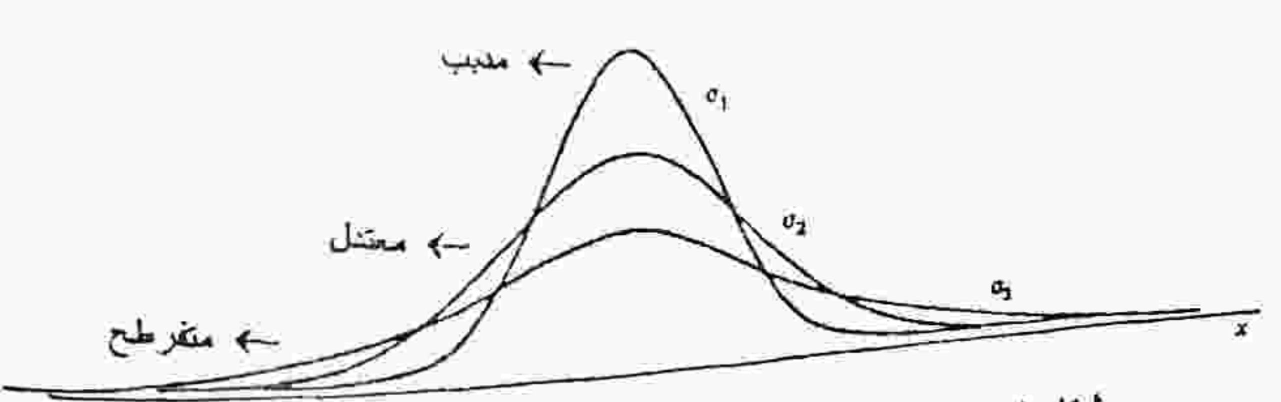
أ- التواء موجب

شكل (14) : الالتواء موجب أو سالب .

7- في بعض الأحيان يستخدم العزم الرابع المركزي حول المتوسط ، μ_4 ، كمقياس لدرجة تدبب أو تفرطح (Kurtosis) منحنى التوزيع الاحتمالي لمتغير عشوائي .ويمكن قياس درجة تفرطح منحنى التوزيع الاحتمالي وفق الصيغة الآتية :

$$\beta = \frac{\mu_4}{\sigma^4} - 3 \qquad (26)$$

وتسمى β بمعامل التفرطح .فإذا كانت $\beta = 0$ فهذا يعنى أن منحنى التوزيع الاحتمالي معتدل التفرطح ، وإذا كانت $\beta > 0$ فهذا يعنى أن منحنى التوزيع الاحتمالي مدبب التفرطح ، وإذا كانت $\beta < 0$ فهذا يعنى أن التوزيع الاحتمالي متفرطح كما هو مبين في شكل (15) .



شكل (15) : منحنيات احتمالية حسب درجة التفرطح .

مثال (27) : إذا كان X متغيراً عشوائياً بدالة كثافة احتمال معرفة كما يلي :

$$f_X(x) = \begin{cases} \lambda\, e^{-\lambda(x-a)} & , a < x < \infty \quad , \lambda > 0 \\ 0 & , ow. \end{cases}$$

اوجد كل من :الوسيط ، الربيع الأدنى ، الربيع الأعلى ، الانحراف الربيعي .

الحـل :

سوف نوجد أولاً دالة التوزيع التراكمي وذلك كما يلي :

إذا كانت $x < a$ فإن $F_X(x) = 0$ وإذا كانت $a \le x < \infty$ فإن

$$F_X(x) = P(X \le x) = \int_a^x f_X(t)\, dt$$

$$= \lambda \int_a^x e^{-\lambda(t-a)}\, dt = 1 - e^{-\lambda(x-a)}$$

وحيث أن $F_X(x)$ دالة متصلة وبالتالي فإنه لأي تجزئ ζ_p حيث أن $0 < p < 1$ فإننا نعلم أن $F_X(\zeta_p) = p$ وعليه فإن

$$F_X(\zeta_p) = 1 - e^{-\lambda(\zeta_p - a)}$$

$$\Rightarrow \qquad \zeta_p = a - \frac{\ln(1-p)}{\lambda}$$

ومنها إذا كانت $p = 0.5$ فإن الوسيط يساوى $\zeta_{0.5} = a + \dfrac{\ln 2}{\lambda}$

وإذا كانت $p = 0.25$ فإن الربيع الأدنى يساوى $\zeta_{0.25} = a + \dfrac{\ln 4 - \ln 3}{\lambda}$

وإذا كانت $p = 0.75$ فإن الربيع الأعلى يساوى $\zeta_{0.75} = a + \dfrac{\ln 4}{\lambda}$

وإن الانحراف الربيعي يساوى $Q = \dfrac{\zeta_{0.75} - \zeta_{0.25}}{\lambda} = \dfrac{\ln 3}{2\lambda}$

علاوة على ذلك حيث أن $0 < \lambda$ و $e^{-\lambda(x-a)} > 0$ وعليه فإن الدالة $f_X(x)$ متناقصة كلما زادت قيمة x ابتداء من a . وبالتالي فإن المنوال ، M ، يساوى $M = a$.

مثال (28) : إذا علمت أن X متغير عشوائي يمثل عدد الزبائن الذين يصلون إلى أحد المصارف في أي دقيقة خلال فترة زمنية مدتها 14 دقيقة والذي دالة كتلة احتماله يمكن كتابتها كالأتي :-

$$p_X(x) = \begin{cases} \dfrac{4^x e^{-4}}{x!} & ,x = 0,1,2,\ldots \\ 0 & ,ow. \end{cases}$$

أوجد كل من :الوسيط ، الربيع الأدنى ، الربيع الأعلى ، الانحراف الربيعي ،ومعاملي الالتواء والتفرطح .

الحل :

حيث أن هذه الدالة تمثل دالة كتلة احتمال لمتغير عشوائي منفصل ،وعليه يمكن وضعها في جدول كما يلي :

$X = x$	$p_X(x)$	$P(X \le x)$	$X = x$	$p_X(x)$	$P(X \le x)$
0	0.0183	0.0183	8	0.0298	0.9786
1	0.0733	0.0916	9	0.0132	0.9918
2	0.1465	0.2381	10	0.0033	0.9971
3	0.1953	0.4334	11	0.0019	0.9991
4	0.1945	0.6288	12	0.0006	0.9997
5	0.1563	0.7851	13	0.0002	0.9999
6	0.1042	0.8893	14	0.0001	1.0
7	0.0595	0.9488			

من تعريف التجزئ ζ_p لدينا $P(X \ge \zeta_p) \ge 1-p$ و $P(X \le \zeta_p) \ge p$ وعليه فإن :

أ- إذا كانت $p = 0.5$ فإننا نلحظ ما يلي :

$$P(X \le 4) \ge 0.5 \quad و \quad P(X \ge 4) \ge 0.5666 \ge 0.5$$

وبذلك تكون $\zeta_{0.5} = 4$ تمثل الوسيط لتوزيع المتغير العشوائي X .

ب- إذا كانت $p = 0.25$ فإن الربيع الأدنى يساوى $\zeta_{0.25} = 3$ وذلك لأن

$$P(X \le 3) \ge 0.25 \quad و \quad P(X \ge 3) \ge 0.7619$$

جـ - إذا كانت $p=0.75$ فإن الربيع الأعلى يساوى $\zeta_{0.75} = 5$ وذلك لأن،

$$P(X \le 5) \ge 0.75 \quad و \quad P(X \ge 5) \ge 0.3712$$

د- الانحراف الربيعي يساوى $Q = \frac{\zeta_{0.75} - \zeta_{0.25}}{2} = \frac{5-3}{2} = 1$

هـ- المنوال هو قيمة X التي تقابل أكبر احتمال وعليه فـإن $M = 3$. أي أن التوزيـع لـه منـوال وحيد .

و- معامل الالتواء (γ) :

حيث أن $\mu = \mu_1 = \lambda = 4$ و $\mu_2 = \lambda = \sigma^2$ و $\mu_3 = \lambda$ و $\mu_4 = 3\lambda^4 + \lambda$

وبالتالي فإن

$$\gamma = \frac{\mu_3}{\sigma^3} = \frac{\mu_3}{(\sigma^2)^{\frac{3}{2}}} = \frac{\lambda}{\lambda\sqrt{\lambda}} = \frac{1}{\sqrt{\lambda}} = \frac{1}{2}$$

ز- معامل التفرطح (β) :

$$\beta = \frac{\mu_4}{\sigma^4} - 3$$

$$= \frac{3\lambda^2 + \lambda}{\lambda^2} - 3 = \frac{1}{4}$$

ويمكن حساب أي تجزئ ذو المرتبة p من خلال الرسم البياني لدالة التوزيع التراكمي ،وذلك على النحو الأتي :ضع قيمة p المرغوب فيها على المحور الرأسي ومنها يتم مد خط مستقيم يوازي المحور الأفقي حتى يلتقي ذلك الخط مع منحنـى دالـة التوزيـع التـراكمي ومـن نقطـة الالتقـاء يتـم إنزال خط عمودي على المحور الأفقي ونقطة التقائهما تمثل قيمة التجزيء المطلوب حسابه .

مثال (29) : إذا كان X متغيراً عشوائياً بدالة كثافة احتمال معرفة كما يلي :

$$f_X(x) = \begin{cases} \frac{3}{8}x^2 & ,0 \le x \le 2 \\ 0 & ,o.w. \end{cases}$$

أوجد :الربيع الأدنى والوسيط والربيع الأعلى بيانياً .

الحـل :

لإيجاد المقاييس المطلوبة بيانياً يجب أولاً إيجاد دالة التوزيع التراكمي وذلك كما يلي:

حيث أنه عندما تكون $x < 0$ تكون $F_X(x) = P(X \le x) = 0$

وعندما $x > 2$ تكون $F_X(x) = 1$

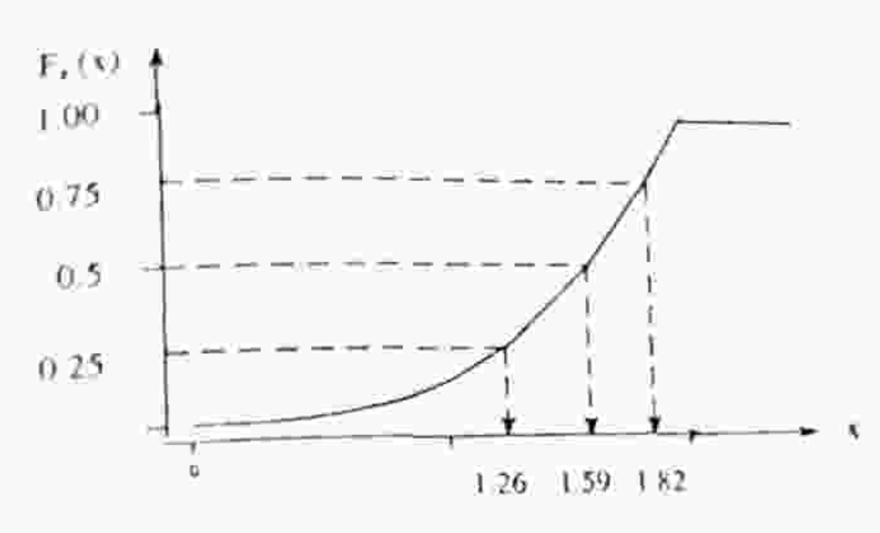
وعندما تكون قيمة x ما بين 0 و 2 تكون دالة التوزيع التراكمي كالأتي :

$$F_X(x) = P(X \le x) = \int_{0}^{x} \frac{3}{8} t^2 \, dt = \frac{x^3}{8}$$

وعليه فإن :

$$F_X(x) = \begin{cases} 0 & , x < 0 \\ \frac{x^3}{8} & , 0 \le x \le 2 \\ 1 & , x > 2 \end{cases}$$

وبتمثيل منحنى دالة التوزيع التراكمي بيانيا كالآتي :



شكل(16) :حساب كل من الوسيط والربيع الأدنى و الربيع الأعلى بيانيا .

ومن الشكل البياني نجد أن $\zeta_{0.75} = 1.82$ و $\zeta_{0.5} = 1.59$ و $\zeta_{0.25} = 1.26$.

مثال (30) : بفرض أن X متغير عشوائي بدالة كتلة احتمال كما يلي :

$$P_X(x) = \begin{cases} \frac{1}{6} & , x = 1,2,3,4,5,6 \\ 0 & , O.W \end{cases}$$

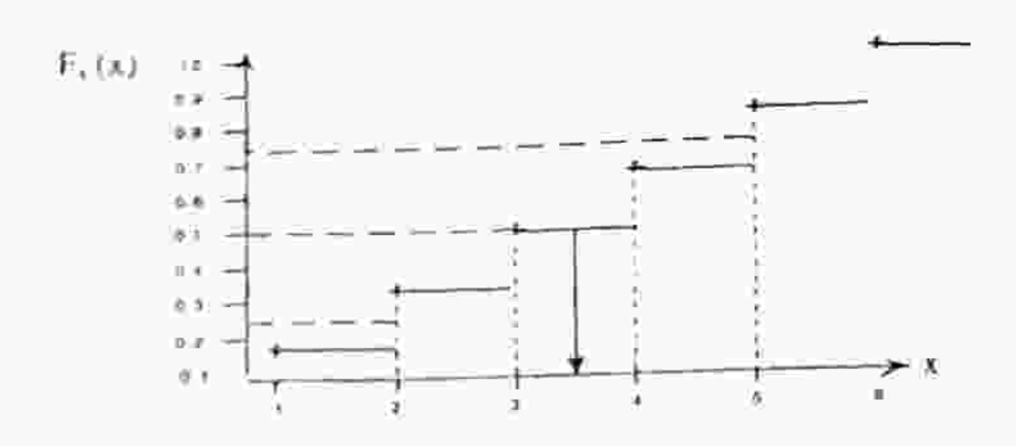
أوجد الوسيط والربيع الأدنى والربيع الأعلى بيانياً .

الحل :

سوف نوجد أولاً دالة التوزيع التراكمي وذلك كالآتي :

$X = x$	1	2	3	4	5	6
$p_X(x) = P(X = x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$
$F_X(x) = P(X \le x)$	$\frac{1}{6}$	$\frac{2}{6}$	$\frac{3}{6}$	$\frac{4}{6}$	$\frac{5}{6}$	$\frac{6}{6}$

وعليه فإن دالة التوزيع التراكمي يمكن تمثيلها بيانيا كما يلي :



شكل (17) : حساب الوسيط والربيع الأدنى والربيع الأعلى بيانيا.

ومن الشكل البياني نجد أن $\zeta_{0.25} = 2$ و $\zeta_{0.5} = \frac{3+4}{2} = 3.5$ و $\zeta_{0.75} = 5$

إذن يتضح مما سبق أنه إذا كان التجريب وحيداً فإن $F_X(\zeta_p) = p$ أما إذا كان هناك أكثر من قيمة واحدة تكون فيها $F_X(\zeta_p) = p$ فإن ζ_p تساوي متوسط أصغر وأكبر القيم التي تكون فيها $F_X(\zeta_p) = p$. أما إذا لا يوجد قيمة تكون فيها $F_X(\zeta_p) = p$ ، نفرض أن x_1 تمثل أكبر قيم x التي تكون فيها $F_X(x) < p$ وأن x_2 تمثل أصغر قيم x التي تكون فيها $F_X(x) > p$ وعليه فإن

$$\zeta_p = \frac{x_2[p - F_X(x_1)] + x_1[F_X(x_2) - p]}{[F_X(x_2) - F_X(x_1)]} \qquad (27)$$

مثال (31) : بفرض أن X متغير عشوائي بدالة توزيع تراكمي كما يلي :

$X = x$	$F_X(x)$	$X = x$	$F_X(x)$
-3.0	0.001	0.5	0.583
-2.5	0.019	1.0	0.692
-2.0	0.057	1.5	0.783
-1.5	0.111	2.0	0.857
-1.0	0.189	2.5	0.954
-0.5	0.298	3.0	1.0
0.0	0.476		

أوجد الوسيط والربيع الأدنى .

الحـل :

الحظ أولاً أن $F_X(0.5) = 0.583 > 0.5$ و $F_X(0.0) = 0.476 < 0.5$

وعليه فإن $x_1 = 0.0$ و $x_2 = 0.5$ ، ومن (27) نجد أن

$$\zeta_{0.5} = \frac{(0.5)[0.5 - 0.476] + (0.0)[0.583 - 0.5]}{0.583 - 0.476} = 0.112$$

ثانياً : حيث أن $F_X(-1.0) = 0.189 < 0.25$ و $F_X(-0.5) = 0.298 > 0.25$

وعليه فإن $x_1 = -1.0$ و $x_2 = -0.5$ وباستخدام (27) نجد أن

$$\zeta_{0.25} = \frac{(-0.5)[0.25 - 0.189] + (-1.0)[0.298 - 0.25]}{0.298 - 0.189} = -0.720$$

وبالمثل يمكن حساب أي تجزؤ نود حسابه.

من الناحية التطبيقية ،يُعد العزم الأول والثاني من أهم العزوم (كما سنرى فيما بعد) ولكن معرفة عزوم ذات مرتبة أعلى نادرًا ما تكون مفيدة ، ولكننا عادة لا نعرف نوعية دالة التوزيع التي سوف تصادفنا في المسائل العملية ،وبالتالي يصبح من الضروري تكوين فكرة حول مقاييس الموقع والتشتت على الأقل .وسوف نعرف الآن نوعا آخر من العزوم وهو العزم العاملي .

تعريف (13) : العزم العاملي Factorial Moment

إذا كان X متغيرًا عشوائيًا بدالة كتلة احتمال $p_X(.)$ أو دالة كثافة احتمال $f_X(.)$ فإن العزم العاملي لهذا المتغير العشوائي ،والذي يرمز له بالرمز $\mu_{[r]}$ ، $(r \geq 1)$ يُعرف كما يلي :

$$\mu_{[r]} = E[(X)_r] = E[X(X-1)\dots(X-r+1)] \quad , r = 1,2,3,\dots$$

1- إذا كان X متغيراً عشوائيا منفصلاً فإن :

$$\mu_{[r]} = E\{\prod_{i=1}^{r}[X-i+1]\} = \sum_{x}\prod_{i=1}^{r}(x-i+1)p_X(x) \qquad (28)$$

2- إذا كان X متغيراً عشوائيا متصلاً فإن :

$$\mu_{[r]} = E\{\prod_{i=1}^{r}[X-i+1]\} = \int_{-\infty}^{\infty}\prod_{i=1}^{r}(x-i+1)f_X(x)\,dx \qquad (29)$$

ولبعض المتغيرات العشوائية (غالبا المنفصلة) فإن العزم العاملي يكون أسهل في حسابه من العزوم السابقة ، وإنه يمكن الحصول على العزم العاملي من العزوم اللامركزية والعكس صحيح ،وذلك على النحو الأتي :

$$r=1 \Rightarrow \mu_{[1]} = E[(X)_1] = E(X) = \mu'_1$$

$$r=2 \Rightarrow \mu_{[2]} = E[(X)_2] = E[X(X-1)] = E(X^2) - E(X) = \mu'_2 - \mu'_1$$

$$r=3 \Rightarrow \mu_{[3]} = E[(X)_3] = E[X(X-1)(X-2)] = \mu'_3 - 3\mu'_2 + 2\mu'_1$$

$$r=4 \Rightarrow \mu_{[4]} = E[(X)_4] = E[X(X-1)(X-2)(X-3)] = \mu'_4 - 6\mu'_3 + 11\mu'_2 - 6\mu'_1$$

$$\vdots \qquad\qquad \vdots$$

وهكذا .

وبنفس الطريقة يمكن الحصول على العزوم اللامركزية من العزم العاملي ، وذلك على النحو الأتي :

$$X = (X)_1 \quad \Rightarrow \quad E(X) = E((X)_1) \quad \Rightarrow \mu'_1 = \mu_{[1]}$$

$$X^2 = (X)_2 + (X)_1 \Rightarrow E(X^2) = E((X)_2) + E((X)_1) \Rightarrow \mu'_2 = \mu_{[2]} + \mu_{[1]}$$

$$X^3 = (X)_3 + 3(X)_2 + (X)_1 \quad \Rightarrow \mu'_3 = \mu_{[3]} + 3\mu_{[2]} + \mu_{[1]}$$

$$\vdots$$

وهكذا .

وكما أشرنا سابقا فإن العزوم تلعب دوراً مهما في الإحصاء من الناحية التطبيقية ، وفي الحقيقة إنه في بعض الأحيان إذا كانت جميع العزوم معروفة فإنه يمكن تحديد دالة كتلة الاحتمال

أو دالة كثافة الاحتمال ، ونظراً لهذه الأهمية فإنه يبدو من المفيد وجود دالة تمثل جميع العزوم ،
إن مثل تلك الدالة تسمى بالدالة المولدة للعزوم .

7 - 3 الدالة المولدة للعزوم Moment Generating Function

تعريف (14) : إذا كان X متغيراً عشوائياً بدالة كتلة احتمال $p_X(.)$ أو دالة كثافة احتمال $f_X(.)$ فإن الدالة المولدة لعزوم هذا المتغير العشوائي ، والتي يرمز لها بالرمز $m_X(t)$ ، تُعرف بأنها $m_X(t) = E(e^{tX})$ ، وعلى أن تكون القيمة المتوقعة موجودة لكل قيم t في الفترة $-s < t < s$ و $s > 0$ وإنه ،

1- إذا كان X متغيراً عشوائياً منفصلاً فإن :

$$m_X(t) = E(e^{tX}) = \sum_x e^{tx} p_X(x) \qquad (30)$$

2- إذا كان X متغيراً عشوائياً متصلاً فإن :

$$m_X(t) = E(e^{tX}) = \int_{-\infty}^{\infty} e^{tx} f_X(x) dx \qquad (31)$$

إن وجود الدالة المولدة للعزوم مرتبط بكون المجموع أو التكامل متقارب على نحو مطلق ، وإذا لم يكن كذلك فعندئذ يقال أن الدالة المولدة للعزوم غير موجودة . وإذا كانت الدالة المولدة للعزوم موجودة فإنه يمكن التعرف على عزوم التوزيع الاحتمالي الذي اشتقت منه الدالة $m_X(t)$ وذلك من خلال تفاضلها وتقييم النتيجة عندما $t = 0$.

الحظ أنه من مفكوك سلسلة مكلورين أن

$$e^{tx} = 1 + tx + \frac{(tx)^2}{2!} + \frac{(tx)^3}{3!} + \frac{(tx)^4}{4!} + \cdots + \frac{(tx)^r}{r!} + \cdots$$

وعليه فإن :

$$E[e^{tX}] = E[1 + tX + \frac{t^2}{2!}X^2 + \ldots + \frac{t^r}{r!}X^r + \ldots]$$

$$= 1 + tE(X) + \frac{t^2}{2!}E(X^2) + \ldots + \frac{t^r}{r!}E(X^r) + \ldots]$$

$$= 1 + t\mu_1' + \frac{t^2}{2!}\mu_2' + \ldots + \frac{t^r}{r!}\mu_r' + \ldots$$

$$= \sum_{i=0}^{\infty} \frac{t^i}{i!}\mu_i'$$

وهكذا نلحظ أن $m_X(t) = E(e^{tX})$ دالة في جميع العزوم حول الأصل μ_i' و $i = 1,2,3,\cdots,r,\cdots$ حيث μ_i' هي معامل $\frac{t^i}{i!}$. وعليه إذا أمكن إيجاد $E(e^{tX})$ فإنه يمكن إيجاد أي عزم من عزوم المتغير العشوائي X ، وحيث أن :

$$m_X(t) = E[e^{tX}] = 1 + t\mu_1' + \frac{t^2}{2!}\mu_2' + \ldots + \frac{t^r}{r!}\mu_r' + \ldots$$

وعليه فإن :

$$m_X'(t) = \frac{d}{dt}E[e^{tX}] = E[\frac{d}{dt}e^{tX}] = \mu_1' + \frac{2t}{2!}\mu_2' + \ldots + \frac{rt^{r-1}}{r!}\mu_r' + \ldots$$

وبوضع $t = 0$ نجد أن :

$$m_X'(0) = E(X) = \mu_1'$$

وبالمثل

$$m_X''(t) = \mu_2' + t\mu_3' + \ldots + \frac{t^{r-2}}{(r-2)!}\mu_r' + \ldots$$

$$\Rightarrow \qquad m_X''(0) = E(X^2) = \mu_2'$$

وبالاستمرار في تفاضل الدالة المولدة للعزوم k من المرات ووضع $t = 0$ نجد أن

$$m_X^k(0) = \left.\frac{d^k m_X(t)}{dt^k}\right|_{t=0} = E(X^k) = \mu_k'$$

حيث افترضنا إمكانية استبدال التفاضل بالتوقع ، أي افترضنا أن

$$\frac{d}{dt}[\sum_x e^{tx} p_X(x)] = \sum_x \frac{d}{dt}[e^{tx} p_X(x)]$$

في حالة ما يكون المتغير العشوائي منفصل . وإن

$$\frac{d}{dt}[\int e^{tx} f_X(x)\,dx] = \int \frac{d}{dt}[e^{tx} f_X(x)]dx$$

وهذا الافتراض من الممكن دائما تبريره وفى الواقع فهو في حالة المتغير العشوائي المتصل .وهذا الافتراض من الممكن دائما تبريره وفى الواقع فهو صحيح لجميع التوزيعات التي سوف ندرسها في هذا الكتاب .

لاحظ أن المعامل $\frac{t^r}{r!}$ في $m_{(X-a)}(t)$ يعطى $\mu_r(a)$ حيث $\mu_r(a) = E[(X-a)^r]$ أي العزم r حول النقطة a . ونلاحظ من ذلك أيضا إن وجود الدالة المولدة للعزوم اللامركزية مرتبط بوجود الدالة المولدة للعزوم المركزية .

مثال (32) : بفرض أن X متغير عشوائي بدالة كثافة احتمال معرفة كما يلي :

$$f_X(x) = \begin{cases} e^{-x} & ,x > 0 \\ 0 & ,ow. \end{cases}$$

أوجد الدالة المولدة لعزوم المتغير العشوائي X ومنها أوجد القيمة المتوقعة والتباين .

الحل :

لأي عدد حقيقي t تكون

$$m_X(t) = E[e^{tX}] = \int_0^\infty e^{tx} f_X(x)\,dx$$

$$= \int_0^\infty e^{tx} e^{-x}\,dx = \int_0^\infty e^{(t-1)x}\,dx = \frac{1}{1-t} \quad , t < 1.$$

وهذا التكامل سيكون متقاربا إذا وفقط إذا كانت $t < 1$. وعليه فإن $m_X(t)$ تكون موجودة عندما $t < 1$. حيث أن $m_X(t)$ موجودة لجميع قيم t في فترة حول النقطة $t = 0$ ، وعليه فإن جميع عزوم المتغير العشوائي X موجودة ، وبالتالي فإن :

$$m'_X(t) = \frac{d\,m_X(t)}{dt} = \frac{d}{dt}(\frac{1}{1-t}) = \frac{1}{(1-t)^2}$$

$$m''_X(t) = \frac{d^2\,m_X(t)}{dt^2} = \frac{d}{dt}[\frac{1}{(1-t)^2}] = \frac{2}{(1-t)^3}$$

وعندما $t = 0$ نجد إن

$$m'_X(0) = 1 \Rightarrow E(X) = 1$$

$$m''_X(0) = 2 \Rightarrow E(X^2) = 2$$

وعليه فإن

$$V(X) = m''_X(0) - (m'_X(0))^2$$
$$= 2 - 1 = 1$$

مثال (33) : إذا افترضنا أن X متغير عشوائي بدالة كتلة احتمال معرفة كما يلي :

$$p_X(x) = \begin{cases} p^x (1-p)^{1-x} & , x = 0,1 \ , 0 < p < 1 \\ 0 & , ow. \end{cases}$$

أوجد الدالة المولدة لعزوم هذا المتغير ومنها أوجد القيمة المتوقعة والتباين .

الحـل :

الدالة المولدة لعزوم هذا المتغير العشوائي ستكون كالآتي :

$$m_X(t) = E(e^{tX}) = \sum_{x=0}^{1} e^{tx} p_X(x)$$
$$= \sum_{x=0}^{1} e^{tx} p^x (1-p)^{1-x} = \sum_{x=0}^{1} (1-p)(\frac{pe^t}{1-p})^x$$
$$= (1-p) + pe^t$$

وعليه فإن :

$$m'_X(t) = \frac{dm_X(t)}{dt} = pe^t \quad \Rightarrow m'_X(0) = E(X) = p$$

$$m''_X(t) = \frac{d^2m(t)}{dt^2} = pe^t \quad \Rightarrow m''_X(0) = E(X^2) = p$$

وبالتالي سيكون :

$$V(X) = m''_X(0) - (m'_X(0))^2$$
$$= p - p^2 = p(1-p)$$

سنعرض الآن بعض النظريات الأساسية المتعلقة بالدالة المولدة للعزوم .

نظرية (6) : إذا كان X متغيرا عشوائيا بدالة مولدة للعزوم $m_X(t)$ ،وكان $Y = aX + b$ حيث أن a و b ثابتان ، وكانت $m_Y(t)$ ترمز للدالة المولدة لعزوم المتغير العشوائي Y فإنه لأي قيمة t حيث $m_X(at)$ تكون موجودة فإن

$$m_Y(t) = e^{bt} m_X(at) \ .$$

البرهان :
من تعريف الدالة المولدة للعزوم ، نجد أن

$$m_Y(t) = E[e^{tY}] = E[e^{t(aX+b)}]$$

$$m_Y(t) = E[e^{atX+bt}] = E[e^{atX}\, e^{bt}]$$

$$= e^{bt}\, E[e^{atX}]$$

$$= e^{bt}\, m_X(at)$$

مثال (34) : في المثال (31) أوجد الدالة المولدة لعزوم المتغير العشوائي Y حيث $Y = 3 - 2X$.

الحــل :

من مثال (31) نجد أن $m_X(t) = \dfrac{1}{1-t}$ حيث $t < 1$ ، وعليه فإن الدالة المولدة لعزوم المتغير العشواني Y تكون موجودة عندما $t > -\dfrac{1}{2}$ ، وباستخدام النظرية (5) عندما $a = -2$ و $b = 3$ نجد أن

$$m_Y(t) = e^{3t}\, m(-2t) = \frac{e^{3t}}{1+2t}$$

إن النظرية الآتية ، تعد من أهم النظريات المتعلقة بالدالة المولدة للعزوم وإن برهانها خارج نطاق هذا الكتاب .

نظرية (7) : الوحدانية Uniqueness

إذا كانت الدالة المولدة لعزوم المتغيرين العشوائيين X و Y متساوية ،أي أنه إذا كانت $m_X(t) = m_Y(t)$ لجميع قيم t في فترة حول النقطة $t = 0$. فإن دالتي التوزيع للمتغيرين العشوائيين X و Y يجب أن يكونا متساويين ،أي أن $F_X(.) = F_Y(.)$.

الحظ أن الدالة المولدة للعزوم تعتبر مقياساً ضعيفاً في بعض الأحيان ،وذلك لاعتماد t على نطاق التوزيع الاحتمالي للمتغير العشوائي تحت الدراسة . فمثلاً ، إذا كانت

$$p_X(x) = \begin{cases} \dfrac{6}{\pi^2 x^2} & , x = 1,2,3,\ldots \\ 0 & , ow. \end{cases}$$

تمثل دالة كتلة احتمال للمتغير العشوائي X فإن الدالة المولدة لعزوم هذا المتغير تكون كالآتي :

$$m_X(t) = E[e^{tX}] = \sum_{x=1}^{\infty} e^{tx} p_X(x)$$

$$= \sum_{x=1}^{\infty} \frac{6e^{tx}}{\pi^2 x^2}$$

ويمكن استخدام اختبار النسبة (ratio test) لإثبات أن هذه السلسلة متباعدة إذا كانت $t > 0$.
وعليه فإن التوزيع الذي دالة كتلة احتماله كما هي معرفة أعلاه لا توجد له دالة مولدة للعزوم .

3 - 8 الدالة المولدة للاحتمال Probability Generating Function

إذا كان X متغيراً عشوائياً منفصلاً بقيم صحيحة غير سالبة وبدالة كتلة احتمال $p_X(x) = P(X = x)$ ، فإنه لأي عدد حقيقي $0 \le t$ ، تُعرف الدالة

$$\varphi_X(t) = E(t^X) = \sum_{k=0}^{\infty} t^k p_X(k) \qquad , |t| \le 1$$

على أنها الدالة المولدة لاحتمال المتغير العشوائي X .

إن للدالة المولدة للاحتمال عدة خصائص مفيدة في تقييم مزايا التوزيع الاحتمالي الذي اشتقت منه ،ومن أهم هذه الخصائص ما يلي :

1- إن الدالة المولدة لاحتمال متغير عشوائي X تحدد دالة كتلة احتماله ، وفي الواقع أن

$$\varphi_X(t) = p_X(0) + \sum_{k=1}^{\infty} t^k p_X(k)$$

وعليه فإن $\varphi_X(0) = p_X(0) = P(X = 0)$ ، وإن

$$p_X(k) = P(X = x) = \frac{1}{k!} \varphi_X^{(k)}(0) \qquad , k = 1, 2, \ldots$$

حيث $\varphi_X^{(k)}(0) = \left. \frac{d^k \varphi_X(t)}{dt^k} \right|_{t=0}$.

2- إذا كان X متغيراً عشوائياً ذا قيم صحيحة وبدالة كتلة احتمال $p_X(.)$ ، وكانت $\varphi_X(t)$ تمثل الدالة المولدة لاحتمال المتغير العشوائي X ،وكان العزم العاملي r ،أي $\mu_{[r]} = E[(X)_r]$ محدود ، فإن :

$$\varphi_X^{(r)}(1) = E[(X)_r]$$
$$= E[X(X-1)(X-2)...(X-r+1)]$$

3- إذا كان a و b عددين صحيحين وغير سالبين ، وكان $Y = aX + b$ فإن

$$\varphi_Y(t) = E[t^Y] = E[t^{aX+b}]$$
$$= t^b\ E[t^{aX}] = t^b\ \varphi_X(t^a)$$

4- يمكن بسهولة إدراك العلاقة بين الدالة المولدة للعزوم والدالة المولدة للاحتمال وذلك كالآتي :

$$m_X(t) = E(e^{tX}) = E(e^t)^X = \varphi_X(e^t)$$

مثال (35) : إذا كانت

$$p_X(x) = P(X = x) = \begin{cases} \dfrac{e^{-3}\,3^x}{x!} & ,x = 0,1,2,... \\ 0 & ,o.w. \end{cases}$$

تمثل دالة كتلة الاحتمال للمتغير العشوائي X . فأوجد

أ- الدالة المولدة لاحتمالات هذا المتغير ومنها أوجد $V(X)$ و $P(X \geq 3)$.

ب- الدالة المولدة لاحتمالات المتغير العشوائي $Y = 3X + 2$.

الحـل :

أ- من تعريف الدالة المولدة للاحتمال نجد أن :

$$\varphi_X(t) = E[t^X] = \sum_{x=0}^{\infty} t^x \frac{e^{-3}\,3^x}{x!}$$
$$= \sum_{x=0}^{\infty} e^{-3}\frac{(3t)^x}{x!} = e^{-3}\sum_{x=0}^{\infty}\frac{(3t)^x}{x!} = e^{-3}e^{3t} = e^{3(t-1)}$$

وعليه من الخاصية (2) نجد أن :

$$\varphi_X^{(1)}(t) = 3e^{3(t-1)} \quad \Rightarrow \quad \varphi_X^{(1)}(1) = 3 = E(X)$$
$$\varphi_X^{(2)}(t) = 9e^{3(t-1)} \quad \Rightarrow \quad \varphi_X^{(2)}(1) = 9 = E[X(X-1)]$$
$$\Rightarrow \quad V(X) = E[X(X-1)] + E(X) - [E(X)]^2$$
$$= 9 + 3 - (3)^2 = 3$$

ومن الخاصية (1) نجد أن

$$p_X(k) = \frac{\varphi_X^{(k)}(t)}{k!} \quad ,k = 0,1,2,\dots$$

$$p_X(0) = \varphi_X(0) = e^{-3} = \frac{3^0 e^{-3}}{0!} \qquad (\because 0! = 1)$$

$$p_X(1) = \frac{\varphi_X^{(1)}(0)}{1!} = \frac{3e^{-3}}{1!}$$

$$p_X(2) = \frac{\varphi_X^{(2)}(0)}{2!} = \frac{9e^{-3}}{2!} = \frac{3^2 e^{-3}}{2!}$$

وعليه فإن الاحتمال المطلوب يكون على النحو الأتي :

$$P(X \geq 3) = 1 - P(X < 3) = 1 - P(X \leq 2)$$
$$= 1 - \{p_X(0) + p_X(1) + p_X(2)\}$$
$$= 1 - \{e^{-3} + 3e^{-3} + \frac{3^2}{2}e^{-3}\} = 1 - (\frac{17}{2})e^{-3} = 0.5768$$

ب- من الخاصية (3) حيث $a = 3$ و $b = 2$ إن الدالة المولدة لاحتمال المتغير العشوائي $Y=3X+2$ تكون كما يلي :

$$\varphi_Y(t) = t^2 \varphi_X(t^3) = t^2 e^{3(t^3-1)}$$

مثال (36) : إذا كان X متغيرا عشوائيا بدالة كتلة احتمال كالأتي :

$$p_X(x) = \begin{cases} \frac{1}{n+1} & ,x = 0,1,2,3,\dots,n \\ 0 & ,o.w. \end{cases}$$

أوجد الدالة المولدة لاحتمال المتغير العشوائي X .

الحـل :

من تعريف الدالة المولدة لاحتمال المتغير العشوائي نجد أن :

$$\varphi_X(t) = E[t^X] = \sum_{x=0}^{n} t^x\, p_X(x)$$

$$= \sum_{x=0}^{n} \frac{t^x}{n+1} = \frac{1 + t^2 + t^3 + \ldots + t^n}{n+1}$$

$$= \begin{cases} \dfrac{(1 - t^{n+1})}{(1-t)(n+1)} & , t \neq 1 \\ 1 & , t = 1 \end{cases}$$

مثال (37) : إذا كان X متغيرا عشوائيا يمثل زمن الانتظار للحصول على أول صورة عن إلقاء عملة مترية تكرارا ، أي أن :

$$P(X=1) = \frac{1}{2}\,,\ P(X=2) = \frac{1}{2^2}\ ,\ldots, P(X=k) = \frac{1}{2^k}\ \cdots$$

فأوجد الدالة المولدة لاحتمال هذا المتغير ومنها أوجد $\mu_{[k]} = E[(X)_k]$.

الحل :

من التعريف نجد أن :

$$\varphi_X(t) = E[t^X] = \sum_{x=1}^{\infty} t^x\, p_X(x) = \sum_{x=1}^{\infty} t^x \frac{1}{2^x} = \sum_{x=1}^{\infty} (\frac{t}{2})^x$$

$$= (\frac{1}{2})t + (\frac{1}{2})^2 t^2 + \ldots + (\frac{1}{2})^k t^k + \ldots = \frac{t}{2-t} \quad , |t| \le 1$$

ومنها نجد أن :

$$\varphi_X^{(1)}(t) = \frac{2}{(2-t)^2} \qquad \Rightarrow \qquad \varphi_X^{(1)}(1) = 2$$

$$\varphi_X^{(2)}(t) = \frac{4}{(2-t)^3} \qquad \Rightarrow \qquad \varphi_X^{(2)}(1) = 4$$

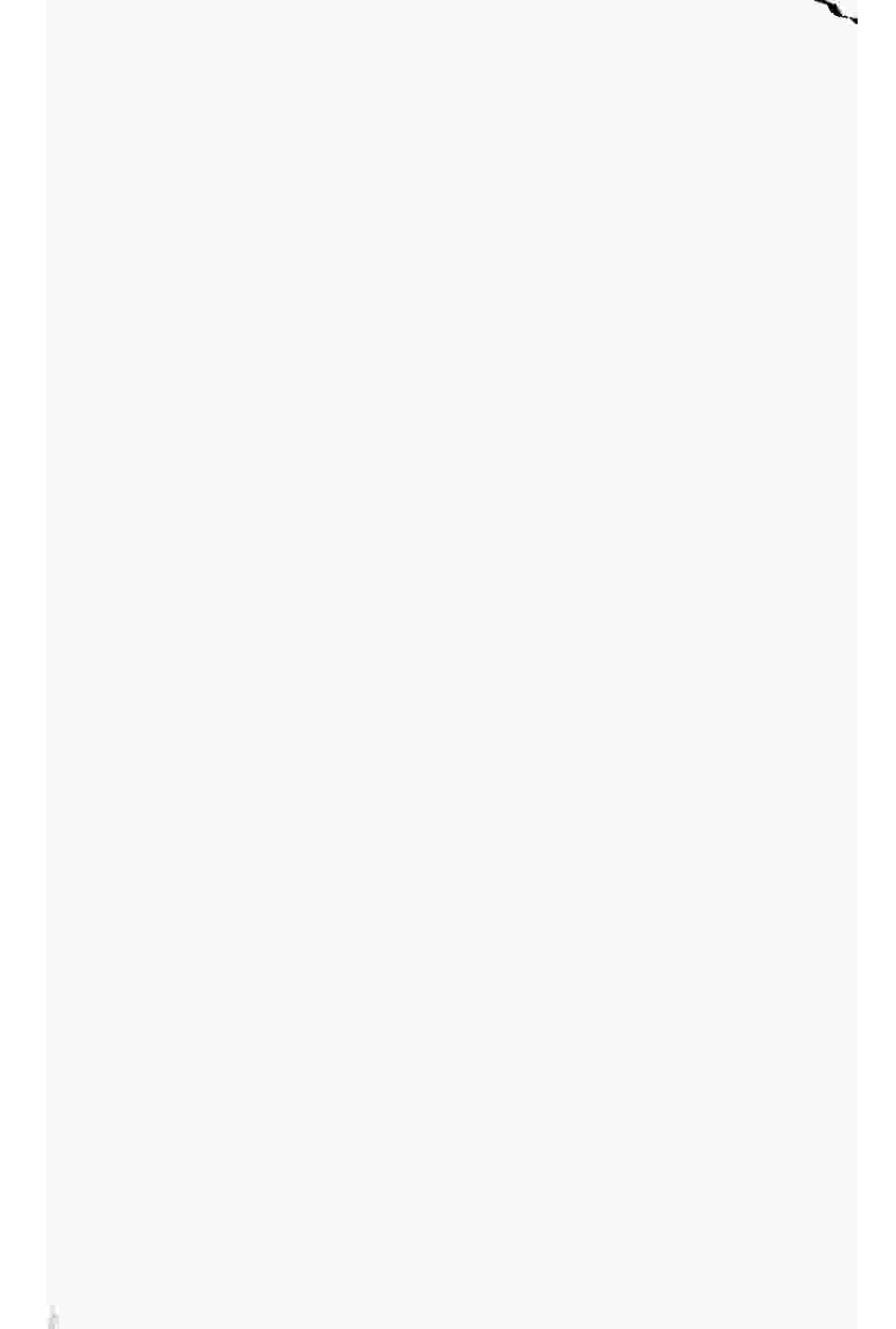
$$\vdots \qquad\qquad\qquad \vdots$$

$$\varphi_X^{(k)}(t) = \frac{2k!}{(2-t)^{k+1}} \qquad \Rightarrow \qquad \varphi_X^{(k)}(1) = 2k!$$

وعليه فإن جميع العزوم متواجدة ، وإن

$$\mu_{[k]} = E[(X)_k] = (K+1)! \qquad , k = 1,2,3,\cdots$$

سوف ننهي هذا البند بملاحظة مهمة جدا وهي أن $\varphi_X(t)$ تكون دالة مولدة للاحتمال إذا وفقط إذا كانت $p_X(k) \ge 0$ حيث $k = 0,1,2$، وأن $\varphi_X(1) = 1$. فمثلاً إذا كان X متغيراً عشوائياً وكانت $\varphi_X(t) = \frac{2}{1+t}$ فإنها لا تمثل دالة مولدة لاحتمالات هذا المتغير العشوائي، وذلك لأن ليست جميع المعاملات غير سالبة، بالرغم من أن $\varphi_X(1) = 1$.



تمرينات Exercises

1- إذا كان X متغيرا عشوائيا يمثل عدد الأبناء الذكور في أسرة لها ثلاثة أطفال تم اختيارها بطريقة عشوائية من مجتمع معين :

ا- ما هي القيم الممكنة للمتغير العشوائي X .

ب- أصف الأحداث الآتية بدلالة المتغير العشوائي X :

1- على الأكثر ولدان 2- ولدان فقط 3- أقل من ولدين 4- على الأكثر ولد

جـ- أوجد دالة كتلة الاحتمال للمتغير X ودالة التوزيع التراكمي ومثلهما بيانيا .

2- ألقيت عملة متزنة بشكل متكرر حتى الحصول على صورة لأول مرة ، فإذا كان المتغير X يمثل عدد المرات المطلوبة :

ا- ما هي القيم الممكنة للمتغير العشوائي X .

ب-اكتب الأحداث $(X=1)$ و $(X=2)$.

جـ- أوجد دالة كتلة الاحتمال للمتغير العشوائي X ودالة التوزيع التراكمي ومثلها بيانيا .

3- بفرض أن المتغير العشوائي X يأخذ القيم $3,1,0,-2$ باحتمالات كما يلي :

$$P(X=-2)=0.4\ ,\ P(X=0)=0.1,\ P(X=1)=0.3,\ P(X=3)=0.2$$

مثل دالة كتلة الاحتمال ودالة التوزيع التراكمي بيانيا .

4- يحتوى صندوق على 10 مصابيح كهربائية من بينها 4 غير صالحة ، فإذا تم اختيار عينة عشوائية حجمها 5 من هذا الصندوق وكان المتغير العشوائي X يمثل عدد المصابيح غير الصالحة بالعينة :

ا- ما هي القيم الممكنة للمتغير العشوائي X .

ب- أصف الأحداث الآتية :-

1- $(X=0)$ 2- $(X\leq 3)$ 3- $(X\geq 4)$

جـ- أوجد دالة كتلة الاحتمال للمتغير العشوائي X ومثلها بيانيا .

د- أوجد دالة التوزيع التراكمي للمتغير العشوائي X ومثلها بيانيا ، ثم أوجد قيمة احتمال الأحداث التي بالفقرة (ب) .

5- بفرض أن X متغير عشوائي بدالة كثافة احتمال كما يلي :

$$f_X(x)=\begin{cases} c\,(1-x^3) & ,0<x<1 \\ 0 & ,ow. \end{cases}$$

ا-أوجد قيمة الثابت c .

ب- أوجد $P(X>\frac{1}{2})$ و $P(\frac{1}{3}<X<\frac{2}{3})$

جـ-مثل دالة كثافة الاحتمال بيانيا .

6- بفرض أن X متغير عشوائي بدالة توزيع تراكمي كما يلي :-

$$F_X(x)=\begin{cases} e^{x-3} & ,x\le 3 \\ 1 & ,x>3 \end{cases}$$

أوجد دالة كثافة الاحتمال ومثلها بيانيا .

7- بفرض أن دالة كثافة احتمال المتغير العشوائي X لها الصيغة الآتية :-

$$f_X(x)=\begin{cases} \frac{x}{8} & ,0\le x\le 4 \\ 0 & ,o.w \end{cases}$$

ا-أوجد قيمة الثابت a بحيث : i- $P(X\le a)=\frac{1}{4}$, ii- $P(X\ge a)=\frac{1}{2}$

ب- أوجد i- $P(X<1)$, ii- $P(1\le X\le 3)$, iii- $P(X\ge 2)$

جـ- مثل دالة كثافة الاحتمال بيانيا ثم أوجد دالة التوزيع ومثلها بيانيا .

8- أي الدوال الآتية تمثل دوال احتمالية مع ذكر السبب :

ا-

$$f_X(x)=\begin{cases} \frac{c}{x} & ,\ 0<x<1 \\ 0 & ,\ ow. \end{cases}$$

ب-
$$f_X(x)=\begin{cases}\frac{c}{1+x} & , x>0\\ 0 & , o.w.\end{cases}$$

جـ-
$$f_X(x)=\begin{cases}\frac{c}{\sqrt{x}} & , 0<x<1\\ 0 & , o.w.\end{cases}$$

د-
$$p_X(x)=\begin{cases}\frac{c}{x} & , x=1,2,3,\ldots\\ 0 & , o.w.\end{cases}$$

هـ-
$$p_X(x)=\begin{cases}\frac{x-6}{5} & , x=7,8,9\\ 0 & , o.w.\end{cases}$$

9- بفرض أن X متغير عشوائي بدالة كتلة احتمال معرفة كما يلي :-

$$f_X(x)=\begin{cases}\frac{c}{x^2} & , x=1,2,\ldots\\ 0 & , o.w.\end{cases}$$

أوجد كلاً من :

ا- قيمة الثابت c .

ب- i- $P(X\le 2)$ ii- $P(X>2)$ iii- $P(2<X<4)$

v- $P(3\le X<5)$ vi- $P(2\le X\le 5)$

vii- $P(X\ge 3)$ iv- $P(3<X\le 5)$

10- أوجد قيمة الثابت c الذى يجعل الدوال الآتية دوال كتلة احتمال :

أ-

$$p_X(x)=\begin{cases}0.2+\dfrac{x}{20} & ,\ x=0,\pm 2\\ c & ,\ x=\pm 1,3\\ 0 & ,\ ow.\end{cases}$$

ب-

$$p_X(x)=\begin{cases}c\,(2x+1) & ,\ x=0,1,2,\ldots,6\\ 0 & ,\ ow.\end{cases}$$

ومنها أوجد دالة التوزيع التراكمي و $P(X<3)$ و $P(2<X<5)$

وقيمة x بحيث $P(X\leq x)<0.5$ و $P(X\leq x)>0.5$

جـ-

$X=x$	-2	-1	0	1	2	4
$p_X(x)=P(X=x)$	0.3	0.1	c	0.2	0.1	c

ومنها أوجد $P(X<0)$ و $P(X=0\mid X<0)$ و $P(X\geq 2\mid X>0)$.

11- أوجد قيمة الثابت c الذي يجعل الدوال الآتية دوال توزيع تراكمية :-

أ-

$$F_X(x)=\begin{cases}0 & ,\ x<0\\ c(x-1)^3 & ,\ 1<x\leq 3\\ 1 & ,\ x\geq 3\end{cases}$$

ب-

$$F_X(x)=\begin{cases}0 & ,\ x<-\theta\\ c\left(\dfrac{x}{\theta}+1\right) & ,\ |x|<\theta\\ 1 & ,\ x>\theta\ ,\ \theta>0\end{cases}$$

12- أي الدوال الآتية تمثل دالة كتلة احتمال ولماذا ؟ وإذا كانت دالة كتلة احتمال فأوجد له التوزيع التراكمي ومثلها بيانياً.

أ- $p_X(x) = \dfrac{x-1}{2^{x-1}} \quad ; \quad x = 1,2,3$

ب- $p_X(x) = \dfrac{x^2-2}{50} \quad ; \quad x = 1,2,3,4,5$

جـ- $p_X(x) = \dfrac{x}{21} \quad ; \quad x = 1,2,3,4,5,6$

د- $p_X(x) = \left(\frac{1}{4}\right)^x \quad ; \quad x = 0,1,2$

13- ألقيت زهرتي نرد متزنتين معاً مرة واحدة .أوجد دالة كتلة الاحتمال ودالة التوزيع التراكمي ومثلهما بيانياً ، ثم أوجد القيمة المتوقعة والتباين في كل حالة من الحالات التالية :

أ- إذا كان المتغير العشوائي X يمثل مجموع العددين الظاهرين على الوجهين العلويين .

ب-إذا كان المتغير العشوائي X يمثل الفرق المطلق بين العددين المذكورين ،

جـ- إذا كان المتغير العشوائي X يمثل العدد الأعظم بين العددين المذكورين .

14- أوجد الثابت c الذي يجعل من الدالة الآتية دالة كتلة احتمال للمتغير العشوائي X ، ثم أوجد دالة التوزيع التراكمي ومثلها بيانياً ، والدالة المولدة لعزوم المتغير العشوائي X ومنها أوجد القيمة المتوقعة والتباين لهذا المتغير حيث $Y = 3X - 4$ ثم أوجد الدالة المولدة لاحتمالات المتغير العشوائي X ومنها أوجد $P(X \leq 2)$. حيث

$$p_X(x) = P(X = x) = c\binom{5}{x} \qquad ; \quad x = 0,1,2,3,4,5$$

15-يحتوي صندوق على كرات متشابهة في الحجم ، 4 بيضاء و 6 حمراء . سحبت عينة من كرتين عشوائياً ، وعلى فرض أن المتغير العشوائي X يمثل عدد الكرات البيضاء في العينة ، أوجد دالة كتلة احتمال المتغير X ، إذا كان السحب : (أ) مع الإعادة ، (ب) بدون إعادة .

16- أوجد قيمة الثابت k الذي يجعل من الجدول الآتي توزيعاً احتمالياً ، ومن ثم أوجد القيمة المتوقعة μ والتباين σ^2 ، وكذلك $P(\mu - 2\sigma < X < \mu + 2\sigma)$.

X	1	2	3	4	5	6	7
$P(X = x)$	1/8	1/16	1/8	1/16	1/8	1/16	k

17- إذا كان X متغيراً عشوائياً دالة كتلة احتماله كالآتي :

$$p(x)=\frac{\left(|x|+1\right)^2}{9} \qquad ; x=-1,0,1$$

أوجد $E(X)$ و $E(X^2)$ و $E(3X^2-2X+4)$.

18- إذا كان X متغيراً عشوائياً دالة توزيعه

$$F_X(x)=\begin{cases}0 & , \quad x<0\\ 0.1 & , \quad 0\le x<1\\ 0.4 & , \quad 1\le x<3\\ 0.8 & , \quad 2\le x<3\\ 1 & , \quad x\ge 3\end{cases}$$

أوجد كلاً مما يأتي :

$P(X=i); i=1,2,3$. $P(X\ge 3)$, $P(X\le 1)$, $P(1<X<3)$

19- إذا كان X متغيراً عشوائياً دالة كتلة احتماله كالآتي ، ما قيمة c ؟ ثم أوجد $F_x\left(\frac{9}{2}\right)$.

$$p_X(x)=c(x-1) \qquad x=1,2,3$$.

20- إذا كان X متغيراً عشوائياً دالة توزيعه التراكمية كما يلي :

$$F_X(x)=\begin{cases}0 & , \quad x<-1\\ \frac{1}{4} & , \quad -1\le x<0\\ \frac{1}{3} & , \quad 0\le x<1\\ \frac{1}{2} & , \quad 1\le x<2\\ 1 & , \quad x\ge 2\end{cases}$$

أوجد كلاً من : $P(X\le 1)$, $P(1<X\le 3)$ ، والقيمة المتوقعة والتباين للمتغير العشوائي Y ، حيث $Y = 2X + 10$.

21- أوجد :

أ- قيمة الثابت k (بدلالة a , b) الذي يجعل من الدالة الآتية دالة كثافة احتمال للمتغير العشوائي X .

ب- القيمة المتوقعة والتباين والدالة المولدة لعزومه . حيث

$$f_X(x) = k \quad , \quad a \le x \le b \, , \, 0 < a < b \ .$$

22- أوجد الثابت k الذي يجعل من الدالة الآتية دالة كثافة احتمال للمتغير العشوائي X ، ثم أوجد القيمة المتوقعة والتباين والدالة المولدة لعزومه ، ودالة التوزيع التراكمي ؟ أوجد قيمة $P(0.5 < x \le 1.5)$ مرة من دالة الاحتمال وأخرى من دالة التوزيع .

$$f_X(x) = \frac{x}{4} + k \quad 0 \le x \le 2 \ .$$

23- إذا كان X متغيراً عشوائياً متوسطه 1.75 ، ودالة كثافة احتماله $f_X(.)$ ، أوجد كلاً من a و b ، ثم التباين إذا علمت أن :

$$f_X(x) = a\,x + b \qquad 0 \le x \le 3 \ .$$

24- أوجد قيمة الثابت c الذي يجعل كلا من الدوال الآتية دالة كثافة احتمال للمتغير العشوائي X :

أ- $f_X(x) = \frac{3}{16} x^2 \ , \ -c < x < c$

ب- $f_X(x) = \frac{c}{x^{3/4}} \ , \ 0 \le x \le 2$

جـ- $f_X(x) = \frac{x^2}{4} \ , \ 0 < x < c$

د- $f_X(x) = \frac{c}{\sqrt{x}} \ , \ 0 \le x \le 1$

هـ- $f_X(x) = c\sqrt{x} \ , \ 0 < x < 4$

و- $f_X(x) = 4x^c \ , \ 0 \le x \le 1$

25 -على فرض أن الدالة الآتية تمثل دالة كثافة احتمال للمتغير العشوائي X ، أوجد القيمة المتوقعة والتباين ، ودالة التوزيع التراكمي .

$$f_X(x) = \frac{1}{2} \sin x \quad ; \quad 0 \le x \le \pi \ .$$

26- أوجد الثابت d الذى يجعل من الدالة الآتية دالة كثافة احتمال للمتغير العشوائى X ، ثم أوجد القيمة المتوقعة، وبين أن تباينه غير موجود .

$$f_X(x) = \frac{d}{x^3} \quad ; \quad 1 \le x \le \infty$$

27- من وعاء يحتوى سبع كرات بيضاء وإحدى عشر كرة حمراء ، سحبت كرة واحدة عشوائيا وكانت $X = 1$ إذا كانت الكرة بيضاء و $X = 0$ إذا كانت الكرة حمراء . فأوجد دالة كتلة احتمال المتغير العشوائي X ثم أوجد القيمة المتوقعة والتباين لهذا المتغير .

28-بفرض أن X متغير عشوائي بدالة كتلة احتمال معرفة كما يلي :

$X = x$	0	1	4	9	16
$p_X(x) = P(X = x)$	0.64	0.25	0.09	0.01	0.01

أوجد كلاً من :-

ا- $E(\sqrt{X})$ و $\sigma_X^2 = V(X)$

ب- $E(X - 2\sqrt{X})$ و $\left[E(X^2 - 4X)\right]^{\frac{1}{2}}$

و- $E\left[X^2 - 2X + 3\sqrt{X} - 4\right]$

جـ- $\xi_{0.5}$ و $\xi_{0.57}$ والمنوال .

29- إذا كان X متغيرا عشوائيا بدالة توزيع تراكمي كما يلي :-

$$F_X(x) = \begin{cases} 0 & , x < 1 \\ \frac{(x-1)^3}{8} & , 1 \le x < 3 \\ 1 & , x > 3 \end{cases}$$

أوجد كل من :

ا- $E(X)$ و σ_X^2 و $E(\frac{1}{X})$ و $\frac{1}{E(X)}$

ب- $\xi_{0.5}$ و $\xi_{0.25}$ و $\xi_{0.75}$والمنوال .

جـ- $P\left[|X-E(X)|\geq 2\sigma\right]$ ثم أوجد حد أعلى لهذا الاحتمال باستخدام متباينة تشيبشيف.

30- أوجد كلاً من $\xi_{0.5}$، المنوال و $E(X)$ و σ_X^2 و $P\left[|X|\geq 3\sigma\right]$ إذا علمت أن :

$X=x$	-1	0	1	2
$p_X(x)$	0.35	0.30	0.20	0.15

ثم أوجد كلاً من :

أ - حد أعلى للاحتمال $P\left[|X|\geq 3\sigma\right]$ باستخدام متباينة تشيبشيف .

ب- μ_2' و μ_3 و $\mu_{[4]}$ و $m_X(t)$.

31- بفرض أن X متغير عشوائي بدالة كثافة احتمال معرفة كما يلي :-

$$f_X(x)=\begin{cases}\dfrac{3x^2}{8} & ,0\leq x\leq 2\\ 0 & ,\text{ow.}\end{cases}$$

أوجد كلاً من :

أ- $E(X)$ و $E(\frac{1}{X})$ و σ_X^2.

ب - $E\left[(X-\mu_X)^3\right]$ و $E\left[(X-\mu_X)^4\right]$ و μ_4' و $\mu_{[3]}$.

جـ - $m_X(t)$.

32 - إذا كان Y متغيراً عشوائياً بمتوسط μ وتباين σ^2 ودالة مولدة للعزوم $m_Y(t)$ أثبت أن

أ - $E(\frac{Y-\mu}{\sigma})=0$.

ب - $E\{e^{t(\frac{Y-\mu}{\sigma})}\}=e^{-\mu t}m_Y(t/\sigma)$.

جـ - $E[(\frac{Y-\mu}{\sigma})^2]=1$.

33- إذا كان X متغيراً عشوائياً بدالة كثافة احتمال معرفة كما يلي :-

$$f_X(x)=\begin{cases}\frac{1}{3} & ,-1<x<2\\ 0 & ,ow.\end{cases}$$

أوجد كلاً من :-

أ - $m_X(t)$ ومنها أوجد μ_1' و μ_2' و σ_X^2.

ب - $E\left(\frac{(X-\mu_X)}{\sigma^3}\right)$ و $E\left(\frac{(X-\mu_X)^4}{4}\right)$.

جـ - $\xi_{0.5}$ و $\xi_{0.90}$ و $\xi_{0.70}$ والمنوال .

د - $P\left[|X-0.5|\geq\frac{3}{4}\sigma_X\right]$ ثم أوجد حد أعلى لهذا الاحتمال باستخدام متباينة تشيبشيف .

34- بفرض أن X متغير عشوائي بدالة كتلة احتمال معرفة كما يلي :

$X=x$	-2	-1	0	1	2
$p_X(x)$	.08	.18	.48	.18	.08

أوجد كلاً من :-

أ - $\xi_{0.30}$ و $\xi_{0.50}$ و $\xi_{0.75}$ و $\xi_{0.85}$ و المنوال .

ب - $P\left[|X-\mu_X|\geq 1.5\,\sigma_X\right]$.

جـ - μ_3' و μ_3 و $\mu_{[3]}$ و $\mu_{[4]}$.

35- إذا كان متوسط درجات الطلبة بأحد مواد قسم الإحصاء يساوى 50 وانحراف معياري يساوى 10 أوجد كلاً من :-

أ -القيم المعيارية المناظرة للدرجات 45 و 50 و 65 و 70 .

ب - الدرجات الحقيقية المناظرة للقيم المعيارية 1.25، 0.75 ، 0 ،-1.5 .

36- بفرض أن X متغير عشوائي حيث $P(X\geq 0)=1$ و $P(X\geq 10)=0.20$ أثبت أن $E(X)\geq 2$.

37- إذا كان X متغيراً عشوائيا بدالة كثافة احتمال معرفة كما يلي :

$$f_X(x)=\begin{cases}\frac{3}{8}(2-x)^2 & ,0\le x\le 2\\ 0 & ,ow.\end{cases}$$

أوجد كلاً من :

أ- μ'_1 و μ'_2 و μ'_3 و μ_2 و $\mu_{[3]}$.

ب- $\xi_{0.50}$ و $\xi_{0.30}$ و $\xi_{0.90}$ و $\xi_{0.75}$ والمنوال .

38- بفرض أن X متغير عشوائي بدالة كتلة احتمال كالآتي :

$X=x$	$\mu-k\sigma$	μ	$\mu+k\sigma$
$p_X(x)$	$\frac{1}{2k^2}$	$1-\frac{1}{k^2}$	$\frac{1}{2k^2}$

حيث $k\ge 1$ ، أثبت أن :

أ - $\mu_X=\mu$ و $\sigma_X^2=\sigma^2$.

ب - $P(X\le\mu-k\sigma$ أو $X\ge\mu+k\sigma)=\frac{1}{k^2}$.

39- في كل حالة من الحالات الآتية : أوجد كل من : μ_3 و μ_4 و β و $\xi_{0.50}$ و $m_X(t)$ والمنوال .

أ - $$f_X(x)=\begin{cases}2x & ,0\le x\le 1\\ 0 & ,ow.\end{cases}$$

ب - $$f_X(x)=\begin{cases}4x & ,0\le x\le 0.5\\ 4(1-x) & ,0.5< x\le 1\\ 0 & ,ow.\end{cases}$$

جـ - $$f_X(x)=\begin{cases}2(1-x) & ,0\le x\le 1\\ 0 & ,ow.\end{cases}$$

... $E(X) = \mu$ حيث μ ...

40- إذا كان X متغيراً عشوائياً حيث ... $P[|X-\mu| \geq 3\sigma_X] = \frac{1}{9}$ حيث X للمتغير العشوائي.

41- بعرض أن X متغير عشوائي يمثل العمر الزمني الذي يعمره مصباح كهربائي ، والذي يمكن تمثيله بدالة كثافة احتمال معرفة كما يلي :-

$$f_X(x) = \begin{cases} \frac{1}{\beta} e^{-x/\beta} & , x > 0 \quad , \beta > 0 \\ 0 & , ow \end{cases}$$

أوجد كلاً من :

أ - $E(X)$ و σ_X^2 و $m_X(t)$.

ب - $P(X > E(X))$.

حـ - $p[|X-\mu| \leq k\sigma]$.ثم أوجد حداً أعلى لهذا أ الاحتمال باستخدام متباينة تشيبشيف.

42- أوجد التجزيء p $(0 < p < 1)$ (إن وجد) والمنوال في الحالات الآتية :-

أ -
$$p_X(x) = P(X = x) = \begin{cases} \frac{1}{N} & , x = 1,2,3,\ldots,N \\ 0 & , ow. \end{cases}$$

ب -
$$f_X(x) = \begin{cases} 2x e^{-x^2} & , 0 \leq x < \infty \\ 0 & , ow. \end{cases}$$

حـ -
$$f_X(x) = \begin{cases} \frac{12}{\alpha^4} (\alpha - x)^2 & , 0 < \alpha < x \\ 0 & , ow. \end{cases}$$

الفصل الخامس

توزيعات خاصة

Special Distributions

5 - 1 مقدمة Introduction

لقد تعرضنا في الفصول السابقة للتوزيعات الاحتمالية بصفة عامة ، وتعرفنا على بعض خصائصها وذلك من خلال إيجاد المتوسط والتباين والدالة المولدة للعزوم . . . الخ . وفي هذا الفصل سوف نعرف ونناقش عدد من التوزيعات الخاصة ، والتي تستخدم بشكل واسع في التطبيقات الإحصائية والاحتمالات . وسوف نصف بشكل موجز كل توزيع نتعرض له ، وندرس بعض خصائصه الأساسية .

5 - 2 التوزيعات المنفصلة (المنقطعة) Discrete Distributions

في هذا البند نناقش عدة توزيعات منفصلة مع اشتقاق كل من المتوسط والتباين والدالة المولدة للعزوم (إن وجدت) ، مع بعض الأمثلة لتجارب عشوائية ، يمكن أن يكون التوزيع قيد النقاش نموذجاً مناسباً لها .

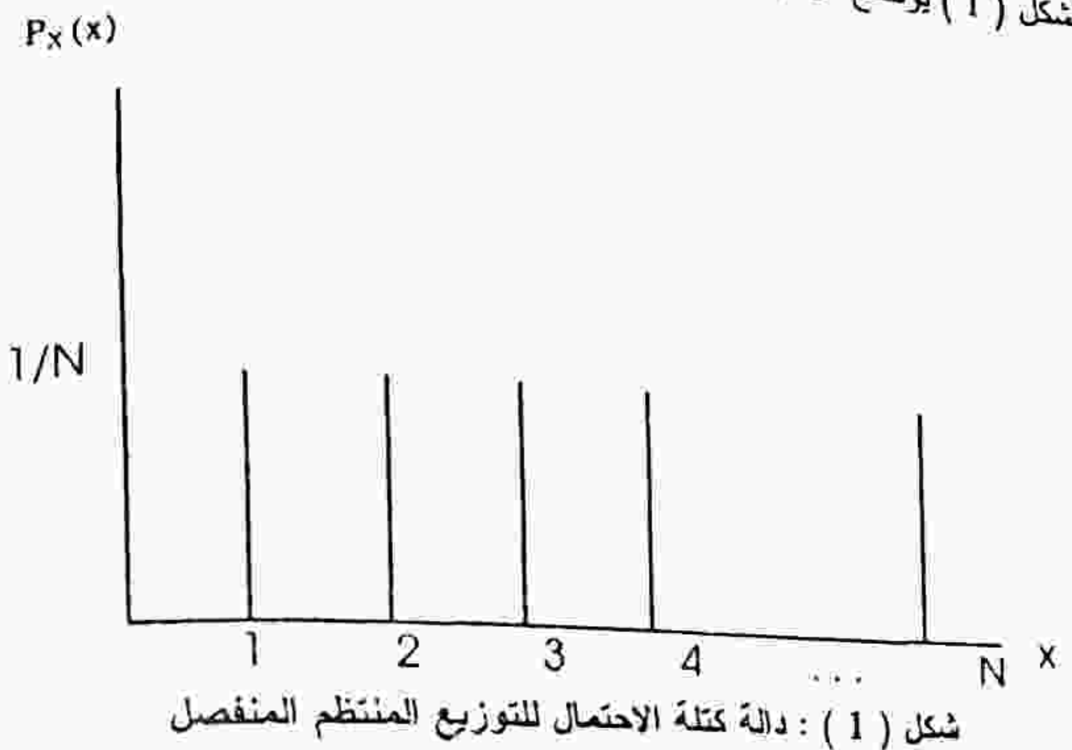
5 - 2 - 1 التوزيع المنتظم المنفصل Discrete Uniform Distribution

يعد التوزيع المنتظم من ابسط أنواع التوزيعات الاحتمالية المنفصلة . حيث يستخدم هذا التوزيع في التجارب التي تتصف نتائجها بأن لها نفس الفرصة في الحدوث . فمثلاً ، عند إلقاء مكعب (زهرة) نرد مرة واحدة وتعريف المتغير العشوائي X بأنه عدد النقاط التي تظهر على الوجه العلوي ، فإن المتغير العشوائي X يتوزع وفق التوزيع المنتظم المنفصل . وعند سحب بطاقة من بين 52 بطاقة والمتغير العشوائي X يمثل نوع البطاقة التي يتم سحبها فإن المتغير X يتوزع وفق التوزيع المنتظم المنفصل . . . الخ . فإذا كان X متغيراً عشوائياً منفصلاً بدالة احتمال معرفة كالآتي :

$$p_X(x) = P(X = x) = \begin{cases} \dfrac{1}{N} & , \quad x = 1, 2, \ \ldots \ , N \\ 0 & , \quad \text{خلاف ذلك} \end{cases} \qquad (1)$$

حيث N عدد صحيح موجب . فإن المتغير العشوائي X يتوزع وفق التوزيع المنفصل بمعلمة N

والشكل (1) يوضح الرسم البياني لهذه الدالة .



شكل (1) : دالة كتلة الاحتمال للتوزيع المنتظم المنفصل

من الواضح أن $p_X(x)$ دالة غير سالبة لكل قيمة من قيم X . وأيضاً

$$\sum_{x=1}^{N} p_X(x) = \sum_{x=1}^{N} \frac{1}{N} = N \cdot \frac{1}{N} = 1$$

إن دالة التوزيع التراكمي (c.d.f) لهذا التوزيع تكون كالآتي :

$$F_X(x) = P(X \le x) = \sum_{k=1}^{x} p_X(x) = \sum_{k=1}^{x} \frac{1}{N} = \frac{x}{N} \qquad , x = 1, 2, \ \ldots \ , N$$

وعليـــه فإن :

$$F_X(x) = \begin{cases} 0 & , \ x < 1 \\ \dfrac{x}{N} & , \ x = 1, 2, \ \ldots \ , N \\ 1 & , \ x \ge N \end{cases} \qquad (2)$$

نظرية (1): إذا كان المتغير العشوائي X يتوزع وفق التوزيع المنتظم المنفصل بمعلمة N فإن:

$$\mu_X = E(X) = \frac{N+1}{2}$$

$$\sigma_X^2 = V(X) = \frac{N^2-1}{12}$$

$$m_X(t) = \frac{e^t\left(e^{Nt}-1\right)}{N\left(e^t-1\right)} \quad ,t>0$$

البرهان :

$$E(X) = \sum_{x=1}^{N} x.\frac{1}{N} = \frac{1}{N}\sum_{x=1}^{N} x = \frac{1}{N}.\frac{N(N+1)}{2} = \frac{N+1}{2} \qquad (3)$$

وبالمثل

$$E(X^2) = \sum_{x=1}^{N} x^2.\frac{1}{N} = \frac{1}{N}\sum_{x=1}^{N} x^2 = \frac{1}{N}\left[\frac{N(N+1)(2N+1)}{6}\right]$$

$$= \frac{(N+1)(2N+1)}{6}$$

وعليه فإن :

$$V(X) = E(X^2) - [E(X)]^2 = \frac{(N+1)(2N+1)}{6} - \left[\frac{N+1}{2}\right]^2$$

$$= \frac{(N+1)(N-1)}{12} = \frac{N^2-1}{12} \qquad (4)$$

وإن الدالة المولدة لعزومه حول نقطة الأصل هي :

$$m_X(t) = E(e^{tx}) = \sum_{x=1}^{N} e^{tx}.\frac{1}{N} = \frac{1}{N}\sum_{x=1}^{N} e^{tx}$$

$$= \frac{1}{N}\sum_{x=1}^{N} y^x \qquad ,y=e^t$$

$$= \frac{1}{N}\left[y+y^2+\ \dots\ +y^N\right]$$

$$= \frac{y}{N}\left[1+y+y^2+\ \dots\ +y^{N-1}\right]$$

وحيث أن المجموع داخل القوس يمثل حدود متوالية هندسية منتهية ، أساسها y وأن مجموعها يساوي :

$$\sum_{i=0}^{N-1} y^i = \frac{1-y^N}{1-y}$$

وعليه فإن :

$$m_X(t) = \frac{y}{N}\left(\frac{1-y^N}{1-y}\right) = \frac{e^t(1-e^{Nt})}{N(1-e^t)}$$

$$= \frac{e^t(e^{Nt}-1)}{N(e^t-1)} \qquad , t \neq 0 \qquad (5)$$

مثال (1) : أوجد التوزيع المنتظم لمجموعة جزئية حجمها ثلاثة أشهر من أشهر السنة .

الحل :

حيث إن عدد أشهر السنة يساوي 12 شهراً ، وعليه فإنه يمكننا اختيار ثلاثة أشهر بشكل عشوائي بطرائق عددها $\binom{12}{3} = 220$ طريقة . وبترقيم هذه المجموعات الجزئية من 1 إلى 220 فإن التوزيع الاحتمالي يكون كالآتي :

$$p_X(x) = \begin{cases} \frac{1}{220} & , x = 1, 2, \ldots, 220 \\ 0 & , \text{خلاف ذلك} \end{cases}$$

وحيث أن كل مجموعة جزئية لها نفس الفرصة في الاختيار ، وعليه فإن احتمال اختيار أي مجموعة جزئية ولتكن المجموعة التي رقمها 90 هو :

$$P(X=90) = \frac{1}{220}$$

وان القيمة المتوقعة هي :

$$E(X) = \frac{220+1}{2} = 110.5$$

والتباين هو :

$$V(X) = \frac{(220)^2-1}{12} = 4033.25$$

أن إحدى الحالات الخاصة للتوزيع المنتظم ، والتي تعتبر مهمة هي الحالة التي تكون فيها $N = 2$ وعندئذ يكون $P(X=0) = P(X=1) = 1/2$ والتي سنناقشها بتفصيل في البند القادم .

حالة خاصة أخرى لهذا التوزيع هي عندما يكون التوزيع المنتظم مرتكزاً في نقطة واحدة ولتكن c، عندئذ تكون دالة الاحتمال كالآتي :

$$p_X(c) = P(X = c) = 1$$

لاحظ أن المتغير العشوائي X الذي يتصف بهذه الخاصية يسمى متغير عشوائي " خامل " (Degenerate r. v.) . وعندما نقول بأن المتغير العشوائي X خامل عند c ، نعني بذلك أن $P(X = c) = 1$.

5 - 2 - 2 توزيع بيرنولي Bernoulli Distribution

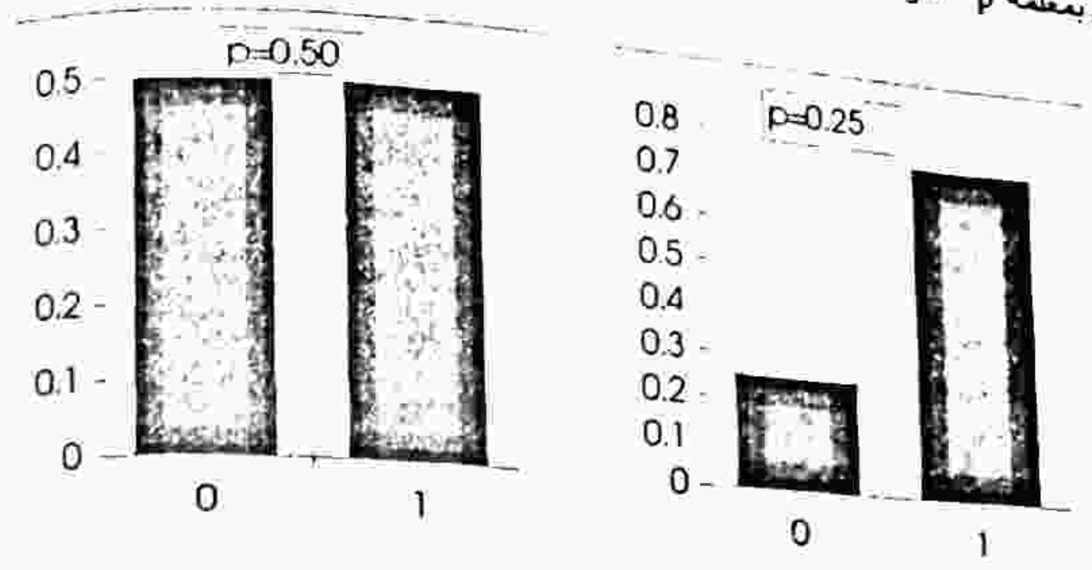
إن أبسط أنواع التجارب هي تلك التي تكون لها نتيجتان ممكنتان ، فمثلاً عند إلقاء قطعة نقود معدنية متزنة مرة واحدة ، فإن النتيجة تكون إما صورة أو كتابة ، أو عند سحب كرة من صندوق فيه m كرة حمراء و n كرة بيضاء ، أو اختيار عنصر من صندوق يحتوى على إنتاج أحد المصانع من سلعة معينة ، بعضها مطابقة للمواصفات ، والبعض الأخر ليست كذلك ، أو اختيار مريض من بين مجموعة من الأشخاص الذين أجريت لهم عمليات جراحية حيث يمكن أن يوجد من بينهم أشخاص كانت عملياتهم ناجحة وبعضهم الأخر فاشلة ، . . . الخ . مما سبق يتضح جلياً أن نتيجة كل تجربة من تجارب بيرنولي تكون أحد ناتجين إما نجاح أو فشل ، فإذا رمزنا لاحتمال النجاح بالحرف p فإن احتمال الفشل هو q حيث $q = 1 - p$. فإذا كان المتغير العشوائي X يمثل نجاح التجربة أو فشلها فإن هذا المتغير يسمى بمتغير بيرنولي ، وتوزيعه الاحتمالي يسمى بتوزيع بيرنولي . ويكون احتمال أن X يساوي قيمة محددة ولتكن x ، حيث $x = 0$ عند فشل التجربة ، و $x = 1$ عند نجاحها هو :

$$p_X(x) = P(X = x) = \begin{cases} p^x q^{1-x} & , \; x = 0, 1 \quad ; \; p + q = 1 \\ 0 & , \; \text{خلاف ذلك} \end{cases} \tag{6}$$

ويطلق على هذه الدالة تسمية دالة الكتلة الاحتمالية (The Probability Mass Function) للمتغير العشوائي X الذي يتوزع وفق توزيع بيرنولي بمعلمة p حيث $0 \le p \le 1$. ويأخذ هذا المتغير قيمتين فقط إما " 0 " أو " 1 " وباحتمالين :

$$P(X = 1) = p \quad , \quad P(X = 0) = 1 - p \tag{7}$$

بيرنولي بمعلمة p . وسوف نرمز لذلك بالرمز X ~ Ber (p) ، ونقرأ أن المتغير العشوائي X يتوزع وفق توزيع بيرنولي بمعلمة p . والشكل (2) يمثل دالة كتلة الاحتمال للمتغير العشوائي X :



شكل (2) : دالة كتلة الاحتمال لتوزيع بيرنولي .

وللتحقق من أن $p_X(x)$ تمثل دالة كتلة احتمال لتوزيع بيرنولي باحتمالات كما هي محددة في (6) ، فيكفي ملاحظة أن :

$$p_X(1) = P(X = 1) = p \quad \text{و} \quad p_X(0) = P(X = 0) = q = 1 - p$$

والجدير بالملاحظة هنا انه عندما $p = q = 1/2$ فإن توزيع بيرنولي في مثل هذه الحالة يمثل توزيعاً منتظماً منفصلاً ، دالة كتلة احتماله تكون على الصورة التالية :

$$p_X(x) = \begin{cases} \frac{1}{2} & , \ x = 0,1 \\ 0 & , \ \text{خلاف ذلك} \end{cases} \qquad (8)$$

ورسمها البياني كما هو موضح في شكل (2)

نظريـــة (2) : إذا كان المتغير العشوائي X يتوزع وفق توزيع بيرنولي بمعلمة p فإن :

$$\mu_X = E(X) = p$$

$$\sigma_X^2 = V(X) = pq$$

$$m_X(t) = q + pe^t$$

البرهان :

من تعريف القيمة المتوقعة نجد أن :

$$E(X)=\sum_{x=0}^{1} x.p_X(x)=\sum_{x=0}^{1} x.P(X=x)$$

$$=\sum_{x=0}^{1} x p^x q^{1-x}=(0)(q)+(1)(p)=p \qquad (9)$$

أي أن المتوسط يساوي p وهو احتمال النجاح في محاولة بيرنولي .

أما التباين فهو :

$$\sigma_X{}^2=V(X)=E(X^2)-[E(X)]^2$$

ولكن :

$$E(X^2)=\sum_{x=0}^{1} x^2 p_X(x)=\sum_{x=0}^{1} x^2 p^x q^{1-x}$$

$$=(0)^2(q)+(1)^2(p)=p$$

وعليه فإن :

$$\sigma_X{}^2=V(X)=p-p^2=p(1-p)=pq \qquad (10)$$

الدالة المولدة لعزوم المتغير العشوائي X هي :

$$m_X(t)=E(e^{tx})=\sum_{x=0}^{1} e^{tx}.p_X(x)$$

$$=\sum_{x=0}^{1} e^{tx} p^x q^{1-x}=\sum_{x=0}^{1}(pe^t)^x q^{1-x}$$

$$=q+pe^t \qquad (11)$$

أما دالة التوزيع التراكمي للمتغير العشوائي X هي ببساطة تكون كالآتي :

$$F_X(x)=\begin{cases}0 & , \quad x<0\\ q & , \quad 0\le x<1\\ 1 & , \quad x\ge 1\end{cases} \qquad (12)$$

وأخيراً إذا كانت المتغيرات العشوائية في المتواليــة اللانهائيــة $X_1, X_2, X_3, \ldots$

مستقلة ومتطابقة التوزيع ، وكل من هذه المتغيرات العشوائية له توزيع بيرنولي بمعلمة p ، فإننا نقول بأنها تشكل متوالية لا نهائية من " محاولات بيرنولي " بمعلمــة p . بالمثل إذا كانت المتغيرات العشوائية $X_1, X_2, \ldots, X_n$ مستقلة ولكل منها توزيع بيرنولي بمعلمة p ، فإننا نقول بأن هذه المتغيرات العشوائية تشكل n من " محاولات بيرنولي " بمعلمة p . فمثلاً ، إذا كان 20 % من الوحدات المنتجة بمصنع معين لسلعة ما معيبة ، وتم اختيار n من هذه الوحدات المنتجة وتم فحصها ، وكان المتغير العشوائي $X_i = 1$ إذا كانت الوحدة i معيبــة و $X_i = 0$ إذا كانت الوحـــدة i ســليمة حيــث $i = 1, 2, \ldots, n$ ، فــإن المتغــيرات العشــوائية $X_1, X_2, \ldots, X_n$ تشكل n من محاولات بيرنولي بمعلمة $p = 1/5$.

5 - 2 - 3 توزيع ذي الحدين Binomial Distribution

يعد هذا التوزيع من أهم التوزيعات الاحتمالية المنفصلة وأكثرها استخداماً ، لما له من تطبيقات عديدة وعلى وجه الخصوص في الرقابة على جودة الإنتاج ، واختبارات النسب . إن أي تجربة عشوائية لها الخواص الآتية تسمى تجربة ذي الحدين :

أ) تتضمن التجربة n من المحاولات .

ب) كل محاولة لها نتيجتان ممكنتان فقط هما " نجاح " أو " فشل " .

جـ) احتمال النجاح ولنكن p ثابت من محاولة إلى أخرى، وعليه فإن احتمال الفشل هو $q = 1 - p$

د) جميع المحاولات مستقلة عن بعضها البعض .

ومن أمثلة تجارب ذي الحدين :

1 . إلقاء قطعة نقود معدنية n من المرات ، حيث المتغير العشوائي X يمثل عدد الصور (أو الكتابات) الممكن الحصول عليها .

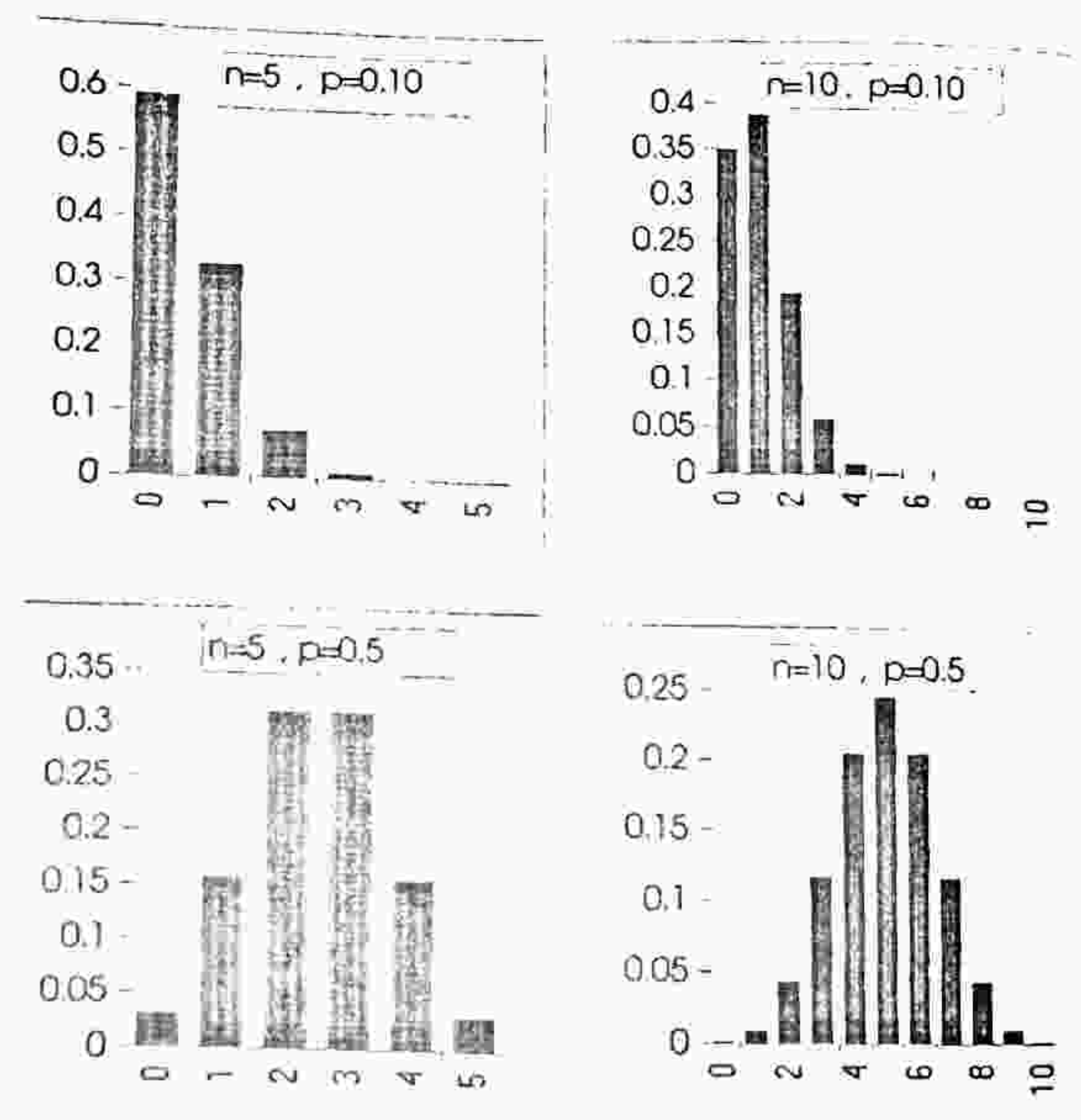
2 . سحب k كرة مع الإعادة من صندوق فيه m كرة بيضاء و n كرة سوداء ، حيث المتغير العشوائي X يمثل عدد الكرات البيضاء الممكن سحبها .

3 . اختيار k عنصر من صندوق يحتوي على m عنصر تالف و n عنصر سليم مع الإحلال ، حيث المتغير العشوائي X يمثل عدد العناصر التالفة .

وبصفة عامة إذا كان المتغير العشوائي X يمثل عدد مرات النجاح في مثل هذا النوع من التجارب ، فإن هذا المتغير يسمى بمتغير ذي الحدين ، وتوزيعه الاحتمالي يسمى بتوزيع ذي الحدين . في مثل هذه الحالة نقول أن المتغير العشوائي X يتوزع وفق توزيع ذي الحدين بمعلمتين n ، p ونرمز لذلك بالرمز $X \sim B(n, p)$ ودالة كتلة احتماله تأخذ الصيغة التالية :

$$p_X(x) = P(X = x) = \begin{cases} \binom{n}{x} p^x q^{n-x} & , x = 0, 1, 2, \ldots, n \\ 0 & \text{خلاف ذلك,} \end{cases} \qquad (13)$$

ويمكن تمثيل دالة الاحتمال بيانياً كما في شكل (3) أدناه :



شكل (3) : دالة كتلة الاحتمال لتوزيع ذي الحدين عند قيم مختلفة للمعلمتين n و p

والجدير بالملاحظة هنا أن تسمية هذا التوزيع بتوزيع ذي الحدين راجع إلى أن الاحتمالات المناظرة للقيم التي من الممكن أن يأخذها المتغير العشوائي X وهي : $x = 0, 1, 2, \ldots, n$ هي الحدود المتتالية في مفكوك ذي الحدين ، أي أن :

$$(p+q)^n = q^n + \binom{n}{1} p q^{n-1} + \binom{n}{2} p^2 q^{n-2} + \ldots + p^n$$

ويلحظ أن الاحتمال $P(X=x)$ دائما موجبا لجميع قيم $x = 0, 1, 2, \ldots, n$ وأن :

$$\sum_{x=0}^{n} p_X(x) = \sum_{x=0}^{n} P(X=x) = \sum_{x=0}^{n} \binom{n}{x} p^x q^{n-x} = (p+q)^n = 1 \quad , p+q=1$$

وهذا يعني أن $p_X(x)$ تفي بشروط دالة كتلة احتمال لمتغير عشوائي منفصل X .

وبالتمعن في دالة كتلة الاحتمال للمتغير العشوائي X (المعادلة رقم (13)) ورسمها البياني (شكل (3)) يمكن استنتاج النقاط التالية :

1 - يتحدد توزيع ذي الحدين بالكامل بمعلومية n و p ، بمعنى انه بمعرفة هاتين القيمتين يمكن معرفة جميع القيم التي من الممكن أن يأخذها المتغير العشوائي X واحتمال كل منها . ويختلف هذا التوزيع عن أي توزيع ذي حدين آخر باختلاف n أو p أو كلاهما .

2 - إذا كانت $p = q = 1/2$ فإن التوزيع متماثل .

3 - إذا كانت $p < q$ يكون التوزيع ملتو إلى اليمين (التواء موجب) .

4 - إذا كانت $p > q$ يكون التوزيع ملتو إلى اليسار (التواء سالب) .

5 - يقترب التوزيع من التماثل كلما كبرت n .

6 - دالة التوزيع التراكمي لهذا التوزيع تكون كالآتي :

$$F_X(x) = P(X \le x) = \sum_{k=0}^{x} \binom{n}{k} p^k q^{n-k}$$

ومنها

$$F_X(x) = (n-x)\binom{n}{x} \int_0^{q} y^{n-x} (1-y)\, dy \qquad (14)$$

إن الطرف الأيمن من المعادلة (14) يسمى تكامل بيتا الناقص (incomplete beta integral) ويمكن التحقق من أن الطرف الأيمن يساوي الطرف الأيسر وذلك من خلال إجراء التكامل بالتجزيء ، وسوف نترك للقارئ التحقق من ذلك .

نظرية (3): إذا كان المتغير العشوائي X يتوزع وفق توزيع ذي الحدين بمعلمتين n و p فإن :

$$m_X = E(X) = n\,p$$

$$\sigma_X^2 = V(X) = n\,p\,q$$

$$m_X(t) = (q + p\,e^t)^n$$

البرهان :

$$E(X) = \sum_{x=0}^{n} x\,p_X(x) = \sum_{x=0}^{n} x\,P(X=x)$$

$$= \sum_{x=0}^{n} x \binom{n}{x} p^x q^{n-x}$$

$$= n\,p \sum_{x=0}^{n} \frac{(n-1)!}{(x-1)!(n-x)!} p^{x-1} q^{n-x}$$

$$= n\,p\,(p+q)^{n-1} = n\,p \qquad (15)$$

ومن تعريف التباين نجد أن :

$$\sigma_X^2 = V(X) = E(X^2) - [E(X)]^2$$

حيث

$$E(X^2) = E[X(X-1)] + E(X)$$

ولكن

$$E[X(X-1)] = \sum_{x=0}^{n} x(x-1) \binom{n}{x} p^x q^{n-x}$$

$$= n(n-1)\,p^2 \sum_{x=2}^{n} \binom{n-2}{x-2} p^{x-2} q^{n-x}$$

$$= n(n-1)\,p^2 (p+q)^{n-2} = n(n-1)\,p^2$$

إذن

$$E(X^2) = E[X(X-1)] + E(X) = n(n-1)\,p^2 + n\,p$$

وعليه فان :

$$\sigma_X^2 = V(X) = n(n-1)\,p^2 + n\,p - n^2 p^2 = n\,p\,(1-p) = n\,p\,q \qquad (16)$$

وأخيراً من تعريف الدالة المولدة للعزوم لمتغير عشوائي منفصل نجد أن :

$$m_X(t) = E(e^{tX}) = \sum_{x=0}^{n} e^{tx}\,p_X(x)$$

$$= \sum_{x=0}^{n} e^{tx} \binom{n}{x} p^x q^{n-x}$$

$$= \sum_{x=0} \binom{n}{x} (pe^t)^x q^{n-x} = (q + pe^t)^n$$

(17)

يمكن النظر لعدد حالات النجاح (X) في المحاولات المستقلة X_i التي عددها n كالآتي :

$$X = \sum_{i=1}^{n} X_i$$

حيث $X_i = 1$ إذا كانت المحاولة i أظهرت نجاحاً ، $X_i = 0$ إذا كانت المحاولة i أظهرت فشلاً. وعليه فإن المتغير العشوائي X يمثل مجموع n من محاولات بيرنولي المستقلة والمتطابقة بمعلمة p . إذاً عندما $n = 1$ فإن توزيع ذي الحدين هو توزيع بيرنولي .

كما أشرنا سابقا يتحدد توزيع ذي الحدين بالكامل بمعلومية n و p . فبمعرفتهما يمكن إيجاد القيمة المتوقعة والتباين وكذلك الدالة المولدة للعزوم ،كما يمكن حساب أي احتمال يتعلق بالمتغير العشوائي X الذي يتوزع وفق توزيع ذي الحدين ، فمثلاً:

$$P(k \le X \le c) = P(X \le c) - P(X \le k-1)$$
$$= \sum_{i=k}^{c} \binom{n}{i} p^i q^{n-i}$$

وأيضاً

$$P(k < X < c) = P(k+1 \le X \le c-1)$$
$$= \sum_{i=k+1}^{c-1} \binom{n}{i} p^i q^{n-i}$$

وفيما يلي بعض الأمثلة التي يمكن من خلالها توضيح بعض التطبيقات لهذا التوزيع .

مثال (2) : إذا علمت أن عشر السيارات التي ينتجها مصنع ما بها خلل ، فإذا اشترى أحد معارض بيع السيارات 4 سيارات فأوجد :

أ – التوزيع الاحتمالي لعدد السيارات التي بها خلل .

ب – بفرض أن المعرض يحقق ربحاً مقداره 200 دينار عن كل سيارة سليمة وخسارة مقدارها 100 دينار عن كل سيارة بها خلل فما هي القيمة المتوقعة للربح أو الخسارة ؟

الحل :

أ – بفرض أن المتغير العشوائي X يمثل عدد السيارات التي بها خلل ، فإن هذا المتغير يتوزع وفق توزيع ذي الحدين بمعلمتين $n = 4$ ، $p = 0.10$ ، أي أن :

$$P(X=x)=\begin{cases}\binom{4}{x}(0.1)^x(0.9)^{4-x} & , x=0,1,2,3,4 \\ 0 & \text{، خلاف ذلك}\end{cases}$$

والجدول التالي يوضح القيم التي من الممكن أن يأخذه المتغير العشوائي X والاحتمالات المناظرة لها :

X = x	0	1	2	3	4
P(X = x)	0.6561	0.2916	0.0486	0.0036	0.0001

ب - بفرض أن المتغير العشوائي Y يمثل مقدار الربح أو الخسارة أي أن $Y = 200(4-X) - 100X$ ، فإن القيم التي من الممكن أن يأخذها هذا المتغير والاحتمالات المناظرة لها تكون على النحو التالي :

Y = y	800	500	200	− 100	− 400
P(Y = y)	0.6561	0.2916	0.0486	0.0036	0.0001

وبذلك تكون القيمة المتوقعة للربح أو الخسارة كالآتي :

$$E(Y)=\sum_{y=0}^{4} y\,P(Y=y)$$

$$=800\times0.6561+500\times0.2416+200\times0.0486+(-100)\times0.0036+(-400)\times0.0001$$

$$=680$$

وعليه فإنه من المتوقع أن يحقق ذلك المعرض ربحاً مقداره 680 دينار وذلك لان إشارة القيمة المتوقعة للربح أو الخسارة موجبة. ألحظ أنه يمكن الحصول على القيمة المتوقعة للمتغير العشوائي Y بدلالة القيمة المتوقعة للمتغير العشوائي X .

مثال (3) : إذا علمت أن احتمال شفاء مريض بالزكام خلال أسبوع دون استخدام الدواء هو 0.45 وعلمت أنه يوجد 8 أشخاص مصابين بالزكام ولم يستخدموا الدواء فأوجد احتمال أن يشفى خلال أسبوع : أ - لا أحد . ب - مريض واحد على الأقل .

جـ- 5 مرضى فقط . د - من 4 إلى 6 مرضى .

الحل :

إذا كان المتغير العشوائي X يمثل عدد المرضى الذين سيتم شفائهم خلال أسبوع دون استخدام الدواء ، فإن هذا المتغير يتوزع وفق توزيع ذي الحدين بمعلمتين n = 8 ، p= 0.45 ، أي أن :

$$P(X=x)=\begin{cases}\binom{8}{x}(0.45)^{x}(0.55)^{8-x} & , x=0,1,2,\ \dots\ ,8\\ 0 & \text{خلاف ذلك ،}\end{cases}$$

وباستخدام جدول توزيع ذي الحدين يتم الحصول على الاحتمالات المطلوبة دون اللجوء إلى استخدام دالة كتلة الاحتمال لتوزيع ذي الحدين ، وذلك على النحو التالي :

أ - باستخدام الجدول (2) حيث n = 8 ، p = 0.45 و x = 0 نجد أن :

$$P(X=0)=0.0084$$

ب - احتمال شفاء مريض واحد على الأقل :

$$P(X\geq 1)=1-P(X=0)=1-0.0084=0.9916$$

جـ - من جدول توزيع ذي الحدي عندما n = 8 ، p = 0.45 و x = 5 نجد أن :

احتمال شفاء 5 مرضى فقط :

$$P(X=5)=0.1719$$

د - احتمال أن يشفى من بينهم من 4 إلى 6 مرضى يمثله الاحتمال التالي : $P(4\leq X\leq 6)$ حيث :

$$\begin{aligned}P(4\leq X\leq 6)&=P(X=4)+P(X=5)+P(X=6)\\&=0.2627+0.1719+0.0703=0.5049\end{aligned}$$

مثال (4) : طائرة تشتغل بأربع محركات مستقلة عن بعضها البعض ، واحتمال توقف أي منها يساوي 0.002 ، ولكي تواصل الطائرة رحلتها يجب أن يشتغل على الأقل اثنان من هذه المحركات ، فإذا قامت الطائرة برحلة جوية فما احتمال أنها ستكمل الرحلة ؟

الحل :

إذن هذه التجربة تتضمن أربعة محاولات مستقلة عن بعضها البعض ، وكل محاولة تتضمن إما المحرك يشتغل (نجاح) أو لا يشتغل (فشل) ، وعليه إذا كان p يمثل احتمال أن

المحرك يشتغل فإن $p = 1 - 0.002 = 0.998$ وهو متساوي لكل محرك . وإن المتغير العشوائي X الذي يمثل عدد المحركات التي تشتغل يتوزع وفق توزيع ذي الحدين بمعلمتين $n = 4$ و $p = 0.998$ وبذلك يكون الاحتمال المطلوب هو :

$$P(X \geq 2) = \sum_{x=2}^{4} \binom{4}{x} (0.998)^x (0.002)^{4-x} = 0.99999997$$

أو باستخدام المكملة :

$$P(X \geq 2) = 1 - P(X < 2) = 1 - [P(X = 0) + P(X = 1)]$$
$$= 1 - \sum_{x=0}^{1} \binom{4}{x} (0.998)^x (0.002)^{4-x}$$
$$= 1 - \left[(0.002)^4 + 4(0.998)(0.002)^3 \right] = 0.99999997$$

مثال (5) : بفرض أن صندوقاً يحتوى على 10 كرات ، منها 4 معيبة . وبفرض أن المتغير العشوائي X يمثل عدد الكرات المعيبة في عينة تتكون من 6 كرات تم سحبها من الصندوق ومع الإعادة . أوجد التوزيع الاحتمالي للمتغير العشوائي X .

الحل :

حيث أن كل سحبة تمثل محاولة من محاولات بيرنولي ، وذلك لأنها من الممكن أن تحتوي على كرات معيبة أو لا تحتوي ، وإن التجربة تتضمن 6 محاولات مستقلة لتجربة بيرنولي والتي فيها $p = 4/10$. وعليه فإن X يتوزع وفق توزيع ذي الحدين بمعلمتين $n = 6$ و $p = 0.4$ ، أي أن :

$$P(X = x) = \begin{cases} \binom{6}{x} (0.4)^x (0.6)^{6-x} & , x = 0, 1, 2, \ldots, 6 \\ 0 & , \text{خلاف ذلك} \end{cases}$$

مثال (6) : إذا علمت أن 10 % من المصابين بمرض معين يتم شفاؤهم ، فإذا تم اختيار عينة عشوائية تتكون من 6 أشخاص يعانون من ذلك المرض . وكان المتغير العشوائي X يمثل عدد الأشخاص الذين سيتم شفاؤهم من هذا المرض ، فأوجد :

أ - $P(X > 1)$, $P(2 < X < 5)$, $P(X = 1)$

ب - القيمة المتوقعة والتباين للمتغير العشوائي X .

جـ - القيمة المتوقعة والتباين للمتغير العشوائي Y حيث $Y=8-5X$.

د - الدالة المولدة للعزوم للمتغير العشوائي X .

الحـل :

أ - من الواضح أن : $X \sim B(n=6, p=0.10)$ وعليه فإن :

$$P(X=1)=\binom{6}{1}(0.10)(0.90)^5=0.354294$$

الحظ أنه بالإمكان إيجاد هذا الاحتمال باستخدام جدول احتمالات توزيع ذي الحدين في آخر هذا الكتاب (جدول رقم (2)) وذلك عندما $n = 6$ و $p = 0.10$ و $x = 1$ نجد أن :

$$P(X=1)=0.3543$$

وأن

$$P(2<X<5)=P(3\le X\le 4)=\sum_{x=3}^{4} P(X=x)$$

$$=\sum_{x=3}^{4}\binom{6}{x}(0.10)^x(0.90)^{6-x}$$

$$=\binom{6}{3}(0.10)^3(0.90)^3+\binom{6}{4}(0.10)^4(0.90)^2$$

$$=0.01458+0.001215=0.015795$$

و

$$P(X>1)=1-P(X\le 1)=1-[P(X=0)+P(X=1)]$$

$$=1-[0.531441+0.354294]=0.114265$$

ب - القيمة المتوقعة للمتغير العشوائي X هي :

$$E(X)=np=6(0.10)=0.6$$

والتباين هو :

$$V(X)=npq=6(0.10)(0.90)=0.54$$

جـ - القيمة المتوقعة للمتغير العشوائي Y :

$$E(Y)=E(8-5X)=8-5E(X)=8-5(0.6)=5$$

و تباينه :

$$\sigma_Y^2=V(Y)=V(8-5X)=V[8+(-5)X]=(-5)^2V(X)$$

$$=25\sigma_X^2=25(0.54)=13.5$$

د- الدالة المولدة للعزوم للمتغير العشوائي X :

$$m_X(t)=\left[q+pe^t\right]^n=\left[0.90+0.10e^t\right]^6$$

مثال (7): إذا كان X متغيراً عشوائياً بدالة مولدة لعزومه كالآتي :

$$m_X(t)=\left(\frac{3}{4}+\frac{e^t}{4}\right)^8$$

فأوجد دالة كتلة الاحتمال لهذا المتغير العشوائي ثم أوجد القيمة المتوقعة والتباين .

الحل :

إنه بمقارنة هذه الدالة بالدالة المولدة لعزوم توزيع ذي الحدين ، ومن خاصية الوحدانية للدالة المولدة للعزوم ، نجد أن X يتوزع وفق توزيع ذي الحدين بمعلمتين $n = 8$ و $p = 1/4$

وبالتالي فإن دالة كتلة الاحتمال لهذا المتغير تكون كالآتي :

$$p_X(x)=P(X=x)=\begin{cases}\binom{8}{x}\left(\frac{1}{4}\right)^x\left(\frac{3}{4}\right)^{8-x} & , x=0,1,2,\ldots,8\\ 0 & \text{خلاف ذلك}\end{cases}$$

وإن

$$E(X) = np = 8\,(1/4) = 2$$

و

$$V(X) = 2n\,p\,q = 8\,(1/4)\,(3/4) = 3/$$

5 - 2 - 4 توزيع ذي الحدين المتعدد The Multinomial Distribution

لنفترض وجود مجتمع يتضمن k من العناصر المختلفة حيث $k \geq 2$ ، وإن نسبة العناصر التي من النوع i في هذا المجتمع هي p_i حيث $i=1,2,\ldots,k$ و $p_i>0$ و $\sum_{i=1}^{k} p_i=1$

بالإضافة إلى ذلك نفترض أن n من العناصر قد تم اختيارها بشكل عشوائي من المجتمع ومع الإعادة ، وبفرض أن X_i ترمز لعدد العناصر المختارة من النوع i $(i = 1, 2, \ldots, k)$ عندئذ يقال بأن المتجه العشوائي $\underline{X}=(X_1,X_2,\ldots,X_k)$ يتوزع وفق توزيع ذي الحدين المتعدد بمعلمتين n و $\underline{p}=(p_1,p_2,\ldots,p_k)$.

يمكننا أن نتصور أن العناصر (n) قد تم اختيارها من المجتمع على أساس عنصر في كل مرة مع الإعادة . وبما أن هذه الاختيارات قد تمت بشكل مستقل عن بعضها البعض ، فإن احتمال أن

يكون العنصر الأول من النوع i_1 هو p_{i_1} واحتمال أن يكون من النوع i_2 هو p_{i_2} ، ....، واحتمال أن يكون من النوع i_n هو p_{i_n} . وعليه فإن احتمال أن تتضمن متوالية النتائج (n) x_1 من العناصر من النوع "1" و x_2 من النوع "2" ، ... ، و x_k من النوع "k" هو $p_1^{x_1} p_2^{x_2} \dots p_k^{x_k}$ مضروباً في العدد الكلي للطرائق المختلفة التي يمكن بها تنظيم أو اختيار العناصر التي عددها n عندما يكون هناك x_i عنصراً من النوع i حيث $i = 1, 2, \dots, k$ هو :

$$\frac{n!}{x_1!\, x_2!\, \dots\, x_k!}$$

وعليه فإن

$$P(X_1 = x_1, X_2 = x_2, \dots, X_k = x_k) = \frac{n!}{x_1!\, x_2!\, \dots\, x_k!} p_1^{x_1} p_2^{x_2} \dots p_k^{x_k} \quad (18)$$

ولأي متجه $\underline{x} = (x_1, x_2, \dots, x_k)$ تُعرف دالة كتلة الاحتمال للمتجه العشوائي $\underline{X} = (X_1, X_2, \dots, X_k)$ بالصيغة الآتية :

$$p_{\underline{X}}(x_1, x_2, \dots, x_k) = P(X_1 = x_1, X_2 = x_2, \dots, X_k = x_k)$$

وإذا كانت $x_1, x_2, \dots, x_n$ تمثل اعداداً صحيحة غير سالبة أي أن $x_i = 0, 1, 2, \dots, n$ و $1 \le i \le k$ بحيث $\sum_{i=1}^{k} x_i = n$ ، فإنه من المعادلة (18) نجد أن

$$p_{\underline{X}}(x_1, x_2, \dots, x_k) = \frac{n!}{x_1!\, x_2! \dots x_k} p_1^{x_1} p_2^{x_2} \dots p_k^{x_k} \quad (19)$$

علاوة على ذلك فإن $p_{\underline{X}}(x_1, x_2, \dots, x_k) = 0$ خلاف ذلك .

نظرية (4) : إذا كانت $X_1, X_2, \dots, X_k$ متغيرات عشوائية لها توزيع ذي الحدين المتعدد بمعلمات n و $p_1, p_2, \dots, p_k$ فإن :

$$E(X_i) = n p_i$$
$$V(X_i) = n p_i (1 - p_i)$$

و $i = 1, 2, \dots, k$

مثال (8) : يفرض أن 23 % من طلبة كلية العلوم في السنة الرابعة و 59 % في السنة الثالثة و 18 % في السنة الثانية ، فإذا تم اختيار عينة عشوائية تتكون من 20 طالباً فما احتمال أن يكون سبعة منهم في السنة الرابعة ؟ وثمانية في السنة الثالثة ؟ وخمسة في السنة الثانية ؟

الحل :

لنفرض أن عدد الطلبة بكلية العلوم كبير بشكل كافي حتى لا يكون الاختيار مع الإعادة أو بدون إعادة ذي أهمية ، وبالتالي يمكننا الافتراض بأن الطلبة تم اختيارهم مع الإعادة .

وبفرض أن المتغير العشوائي X_i يرمز لعدد طلاب السنة i حيث i = 2 ، 3 ، 4 . فإن :

$$p_2 = 0.18 \quad , \quad p_3 = 0.59 \quad , \quad p_4 = 0.23$$

وعليه من المعادلة (19) نجد أن

$$P(X_2 = 5, X_3 = 8, X_4 = 7) = \frac{20!}{5!\,8!\,7!}(0.18)^5(0.59)^8(0.23)^7 = 0.009$$

مثال (9) : إذا كان من المعروف أن أربعة أنواع من معجون الأسنان تباع في السوق بنسبة 40 % ، 15 % ، 20 % ، 25 % على التوالي ، فما احتمال أن يكون في عينة تتكون من 20 مستهلكاً ستة منهم اشتروا النوع الأول ، الربعة اشتروا النوع الثاني ، وخمسة اشتروا النوع الثالث ، وخمسة اشتروا النوع الرابع ؟

الحل :

لنفرض أن المستهلكين يشترون بشكل مستقل ، وإن المتغير العشوائي X_i يمثل عدد الزبائن الذين يشترون المعجون الذي من النوع i من بين العشرون مستهلكاً الذين تم اختيارهم ، وعليه فإن :

$$p_1 = 0.40 \quad , \quad p_2 = 0.15 \quad , \quad p_3 = 0.20 \quad , \quad p_4 = 0.25$$

وبالتالي من المعادلة (19) نجد أن

$$P(X_1 = 6, X_2 = 4, X_3 = 5, X_4 = 5) = \frac{20!}{6!\,4!\,5!\,5!}(0.40)^6(0.15)^4(0.20)^5(0.25)^5$$
$$= 0.007$$

ويمكن حساب العدد المتوقع من المستهلكين الذين اشتروا النوعين الثاني والرابع مثلاً بالعينة وذلك كما يلي :

$$E(X_2) = np_2 = (20)(0.15) = 3$$
$$E(X_4) = np_4 = (20)(0.25) = 5$$

5 - 2 - 5 التوزيع فوق الهندسي The Hypergeometric Distribution

لنفرض وجود مجتمع محدود يتضمن نوعين من العناصر فقط ، وإن عينة بحجم ثابت قد تم اختيارها بشكل متتال ، وبدون إعادة ، وإن الهدف من هذه التجربة هو معرفة عدد عناصر أحد النوعين بالعينة . فمثلاً ، إذا كان لدينا صندوق يحتوي على 4 كرات بيضاء و 5 كرات حمراء وإنه تم سحب ثلاثة كرات وبدون إعادة من ذلك الصندوق ، وكان المتغير العشوائي X يمثل عدد الكرات الحمراء المسحوبة ، فإن المتغير العشوائي X يتوزع وفق التوزيع فوق الهندسي . أو مثلاً ، من الممكن وجود مجموعة من الأشخاص تتكون من 20 رجلاً و 10 نساء ، اختيرت عينة منها حجمها 15 شخصاً وبدون إعادة ، وكان المتغير العشوائي X يمثل عدد النساء بالعينة المختارة فإن X يتوزع وفق التوزيع فوق الهندسي . ولنوضع مثل هذه النوعية من التجارب في صورة دالة احتمالية لنفرض أن صندوقاً يحتوي على N نضيدة جافة ، منها M نضيدة معيبة و N - M نضيدة سليمة . فإذا تم اختيار عينة عشوائية حجمها k من ذلك الصندوق وبدون إعادة ، وكان المتغير العشوائي X يمثل عدد النضائد المعيبة بالعينة التي تم اختيارها فإن هناك :

$\binom{N}{k}$ من الطرائق التي يمكن بها اختيار k نضيدة من بين N نضيدة .

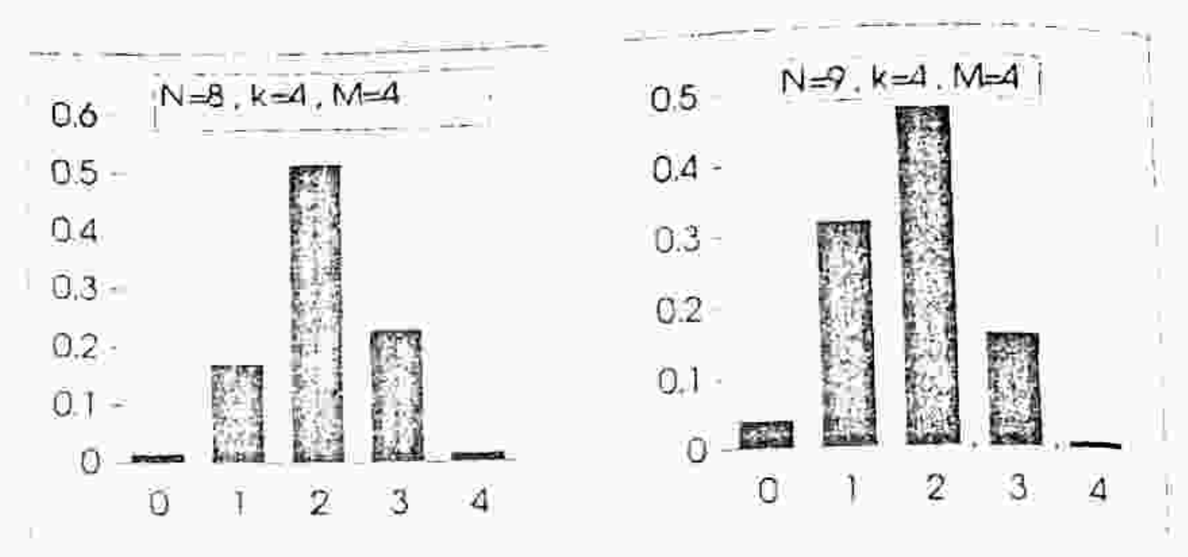
$\binom{M}{x}$ من الطرائق التي يمكن بها اختيار x نضيدة معيبة من بين M نضيدة معيبة.

$\binom{N-M}{k-x}$ من الطرائق التي يمكن بها اختيار (k-x) نضيدة سليمة من بين (N - M) نضيدة سليمة.

وبتطبيق قاعدة الضرب وتعريف الاحتمال (في الفصل الثاني) نجد أن دالة كتلة احتمال المتغير بصفة عامة تكون كالآتي:

$$f_X(x) = P(X = x) = \begin{cases} \dfrac{\binom{M}{x}\binom{N-M}{k-x}}{\binom{N}{k}} & , x = 0, 1, 2, \ldots, k \\ 0 & , \text{خلاف ذلك} \end{cases} \qquad (20)$$

هذه الدالة يطلق عليها تسمية دالة كتلة احتمال التوزيع فوق الهندسي . نقول بان المتغير العشوائي X يتوزع وفق التوزيع فوق الهندسي ، ويرمز لذلك بالرمز : $X \sim H(k, M, N)$. والشكل (4) يبين التمثيل البياني لدالة كتلة الاحتمال للتوزيع فوق الهندسي بمعلمات مختلفة :



شــــكل (4) : دوال كتلة احتمال التوزيع فوق الهندسي

لاحظ أن :

$$\sum_{i=0}^{n} \binom{a}{i}\binom{b}{n-i} = \binom{a+b}{n}$$

وعليه فإن

$$\sum_{x=0}^{k} p_X(x) = \sum_{x=0}^{k} P(X=x) = \sum_{x=0}^{k} \frac{\binom{M}{x}\binom{N-M}{k-x}}{\binom{N}{k}} = 1$$

وبالتالي فإن $p_X(x)$ تمثل دالة كتلة احتمال .

مما سبق يتضح لنا أن تجربة التوزيع فوق الهندسي تشبه إلى حد كبير تجربة ذي الحدين ماعدا أن المعاينة هنا تتم من مجتمع محدود وبدون إعادة (إرجاع) ، وبالتالي فإن المحاولات غير مستقلة .

نظرية (5) : إذا كان المتغير العشوائي X يتوزع وفق التوزيع فوق الهندسي بمعلمات k ، M ، N أي أن $X \sim H(k, M, N)$ فإن القيمة المتوقعة والتباين لهذا التوزيع على الترتيب هما :

$$\mu_x = E(X) = \frac{kM}{N}$$

$$\sigma_x^2 = V(X) = \left(\frac{N-k}{N-1}\right) \cdot k \cdot \frac{M}{N} \cdot \left(1 - \frac{M}{N}\right)$$

البرهان :

من تعريف القيمة المتوقعة نجد أن :

$$\mu_x = E(X) = \sum_{x=0}^{k} x \; p_X(x)$$

$$= \sum_{x=0}^{k} x \frac{\binom{M}{x}\binom{N-M}{k-x}}{\binom{N}{k}}$$

$$= \frac{M}{\binom{N}{k}} \sum_{x=1}^{k} \binom{M-1}{x-1}\binom{N-M}{k-x}$$

$$= \frac{M}{\binom{N}{k}} \sum_{y=0}^{n} \binom{a}{y}\binom{N-a-1}{n-y} \quad , y = x-1 \, , n = k-1 \, , a = M-1$$

$$= \frac{M}{\binom{N}{k}} \cdot \binom{N-1}{n} = \frac{kM}{N} \qquad (21)$$

وبالمثل من تعريف التباين نلحظ أن : $\sigma_x^2 = V(X) = E(X^2) - [E(X)]^2$

ولكن $E(X^2) = E[X(X-1)] + E(X)$ حيث

$$E[X(X-1)]=\sum_{x=1}^{k} x(x-1)\frac{\binom{M}{x}\binom{N-M}{k-x}}{\binom{N}{k}}$$

$$=\frac{M(M-1)}{\binom{N}{k}}\sum_{x=2}^{k}\binom{M-2}{x-2}\binom{N-M}{k-x}$$

$$=\frac{M(M-1)}{\binom{N}{k}}\binom{N-2}{k-2}$$

$$=\frac{M(M-1)k(k-1)}{N(N-1)}$$

ومنها يتضح أن :

$$\sigma_X^2=V(X)=E(X^2)-(E(X))^2$$

$$=\frac{M(M-1)k(k-1)}{N(N-1)}+\frac{kM}{N}-\frac{k^2M^2}{N^2}$$

$$=\frac{kM(N-M)(N-k)}{N^2(N-1)}$$

$$=\left(\frac{N-k}{N-1}\right).k.\frac{M}{N}\left(1-\frac{M}{N}\right) \tag{22}$$

وإذا كانت $M \to \infty$ و $N \to \infty$ بحيث $\frac{M}{N} \to p$ و $0 < p < 1$ فإن :

$$E(X) \to kp \qquad \text{و} \qquad V(X) \to kp(1-p)$$

بشرط $\frac{k}{N} \to 0$. وهذه القيم تمثل المتوسط والتباين لتوزيع ذي الحدين . ومتى كان $\frac{k}{N}$ ليس صغيراً فإن :

$$V(X)=\frac{N-k}{N-1}.kp(1-p)$$

إن المعامل $\frac{N-k}{N-1}$ يسمى معامل التصحيح للتباين عندما تكون المعاينة بدون إعادة وإذا كانت N كبيرة مقارنة بقيمة k فإن هذا المعامل يساوي واحد تقريباً وفي مثل هذه الحالة فإن المعاينة بدون إعادة أو مع الإعادة لا فرق بينهما .

5 - 2 - 6 تقريب التوزيع فوق الهندسي بتوزيع ذي الحدين
Approximating Hypergeometric by Binomial

كما أشرنا أعلاه إنه إذا كانت N كبيرة بالمقارنة مع k فإنه لا يوجد فرق ما بين المعاينة بدون إعادة أو مع الإعادة ، وفي الواقع إنه إذا كانت $N \to \infty$ و $M \to \infty$ بحيث $\frac{M}{N} \to p$ و $0 < p < 1$ فإن :

$$\frac{\binom{M}{x}\binom{N-M}{k-x}}{\binom{N}{k}} \to \binom{k}{x} p^x (1-p)^{k-x}$$

بشرط أن تكون k صغيرة مقارنة مع N . وعليه فإن :

$$P(X=x) = \binom{k}{x} p^x (1-p)^{k-x}$$

لقد وجد أن هذا التقريب جيد عندما $N > 50$ و $\frac{k}{N} \le 0.10$

مثال (10) : يحتوى صندوق على ثلاثة اجهزة كهربائية عاطلة وسبعة صالحة ، فإذا تم اختيار ثلاثة أجهزة وبدون إعادة وكان المتغير العشوائي X يمثل عدد الاجهزة العاطلة بالعينة المختارة
أ - اكتب القيم التى من الممكن ان يأخذها هذا المتغير العشوائى واحتمال كل منها .
ب - اوجد قيمة الاحتمالات التالية :

1 - $P(1 < X \le 3)$ 2 - $P(0 \le X \le 2)$
3 - $P(2 \le X < 3)$ 4 - $P(X = 4)$

الحل :

نلاحظ انه على ضوء الصيغة (20) نجد أن $N = 10$ و $M = 3$ و $k = 3$ و $x = 0 , 1 , 2 , 3$ أى أن :

وعليه فإن دالة كتلة الاحتمال لهذا المتغير العشوائى تكون على النحو التالي :

$$P(X=x)=\begin{cases}\dfrac{\binom{3}{x}\binom{7}{3-x}}{\binom{10}{3}} & , \quad x=0,1,2,3 \\ 0 & , \text{ خلاف ذلك} \end{cases}$$

ومنها يتضح جلياً أن :

$$P(X=0)=\frac{\binom{3}{0}\binom{7}{3}}{\binom{10}{3}}=\frac{35}{120} \quad ، \quad P(X=1)=\frac{\binom{3}{1}\binom{7}{2}}{\binom{10}{3}}=\frac{63}{120}$$

$$P(X=3)=\frac{\binom{3}{3}\binom{7}{0}}{\binom{10}{3}}=\frac{1}{120} \quad ، \quad P(X=3)=\frac{\binom{3}{3}\binom{7}{0}}{\binom{10}{3}}=\frac{1}{120}$$

ويمكن كتابة القيم التى من الممكن ان يأخذه المتغير العشوائى X والاحتمالات المناظرة لها في جدول كما يلي :

$X=x$	0	1	2	3
$P(X=x)$	$\frac{35}{120}$	$\frac{63}{120}$	$\frac{21}{120}$	$\frac{1}{120}$

ب -

-1

$$P(1<X\le 3)=P(X=2)+P(X=3)$$
$$=\frac{21}{120}+\frac{1}{120}=\frac{22}{120}=0.1833$$

- 2

$$P(0\le X\le 2)=P(X=0)+P(X=1)+P(X=2)$$
$$=\frac{35}{120}+\frac{63}{120}+\frac{21}{120}=\frac{119}{120}=0.9917$$

$$P(2 \le X < 3) = P(X = 2) = \frac{21}{120} = 0.175$$

3 -
$$P(X = 4) = 0$$

4 - (لأن X تأخذ القيم 0 ، 1 ، 2 ، 3 فقط)

مثال (11) : فى المثال رقم (7) أوجد القيمة المتوقعة والتباين للمتغير العشوائى X .

الحل :

القيمة المتوقعة والتباين للمتغير العشوائى X تكونا على الترتيب كما يلي :

$$\mu_X = E(X) = \frac{kM}{N}$$

$$= \frac{3(3)}{10} = \frac{9}{10} = 0.90$$

$$\sigma_X^2 = \frac{N-k}{N-1} \cdot k \cdot \frac{M}{N} \cdot \left(1 - \frac{M}{N}\right)$$

$$= \left(\frac{10-3}{10-1}\right)(3)\left(\frac{3}{10}\right)\left(1 - \frac{3}{10}\right) = \frac{49}{100} = 0.49$$

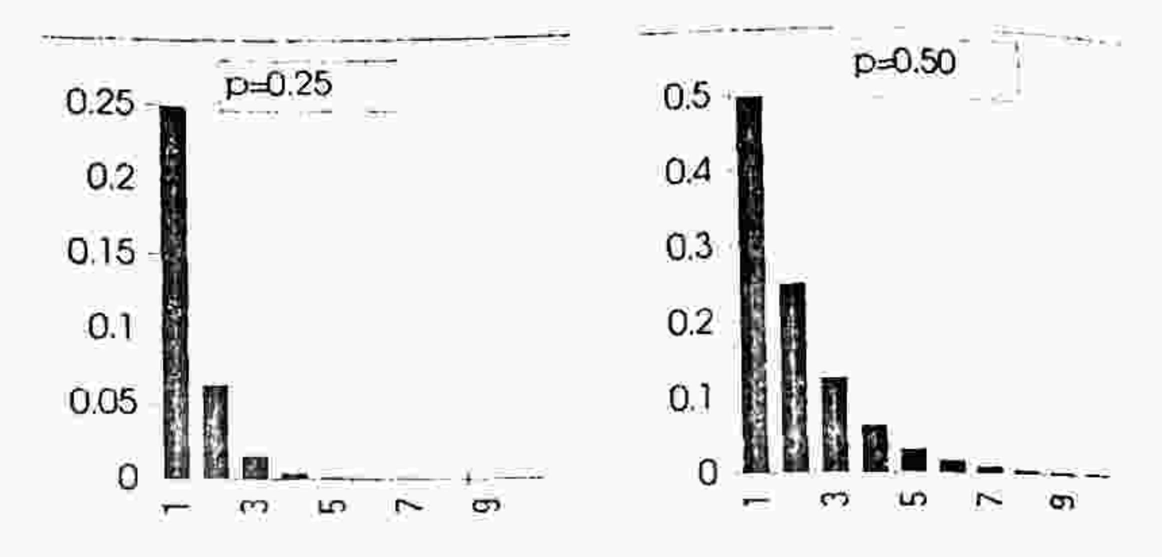
5 ـ 2 ـ 7 التوزيع الهندسي Geometric Distribution

يعد هذا التوزيع من التوزيعات المهمة في التطبيقات الاحصائيـة خاصـة المتعلقة بدراسة الاحصاء السكاني ، حيث يستخدم في دراسة معدلات النمو ومعدلات الوفيات والـولادة ، ...الخ. فإذا كان هناك تجربة احصائية تتألف من متوالية من محاولات بيرنولي المستقلة وكانت نتيجة كل محاولة من هذه المحاولات إما نجاح أو فشل ، وستجرى التجربة لحين الحصـول علـى أول حالة نجاح ، وعلى فـرض أن احتمـال النجـاح ثـابت فـي كـل محاولـة وليكـن p (الحظ هنـا أن عدد المرات (n) متغير عشوائي بينما في حالة ذي الحدين ، عـدد المـرات ثـابت) ، وكـان المتغير العشوائي X يمثل عدد المحاولات المطلوبة لوقوع اول حالة نجاح فإن X يتـوزع وفـق التوزيع الهندسي ،ومن الامثلة على ذلك : إلقاء عملة نقدية معدنية تكراراً حتى ظهور أول صـورة ، إلقاء مكعب نرد حتى ظهـور الرقم " 4 " ، سحب عناصر من صندوق بـه عنـاصر فاسـدة وأخرى صالحة بشكل متتالي ومع الاعادة حتى الحصول على عنصر فاسد ، ... الخ . ويمكـن صياغة تعريف هذا التوزيع كالآتي :

إذا كان X متغيراً عشوائياً يمثل عدد المحاولات المطلوبة للحصول على اول حالـة نجـاح في متوالية من المحاولات المستقلة لتجربة ما وكان احتمال النجاح هو p وهو ثـابت مـن محاولـة لأخرى ، فإن X يتوزع وفق التوزيع الهندسي بدالة كتلة إحتمال معرفة بالصيغة الآتية :

$$P(X=x)=\begin{cases} pq^{x-1} & , x=1,2,\ldots \text{ و } 0<p<1 \\ 0 & , \text{خلاف ذلك} \end{cases} \tag{23}$$

ويرمز لذلك بالرمز $X \sim Ge(p)$. والشكل (5) يبين دالة كتلة الاحتمال .



شـكل (5) : دالة كتلة الاحتمال للتوزيع الهندسي

سنبين اولاً أن (23) تمثل دالة كتلة احتمال ، وفي الواقع إن $P(X=x)\geq 0$ وإن

$$\sum_{x=1}^{\infty} P(X=x)=\sum_{x=1}^{\infty} pq^{x-1}=p\sum_{x=0}^{\infty} q^{x}=p\cdot\frac{1}{1-q}=\frac{p}{p}=1$$

حيث استخدمنا النتيجة الآتية : إذا كانت $0<|\alpha|<1$ فإن :

$$\sum_{x=0}^{\infty} \alpha^{x}=\frac{1}{1-\alpha} \tag{24}$$

وحيث اننا سنوجد كل من المتوسط والتباين فإن العلاقات الآتية مفيدة :

إذا كانت $0<\alpha<1$ فإن :

$$\sum_{x=1}^{\infty} x\alpha^{x-1}=\frac{1}{(1-\alpha)^{2}} \tag{25}$$

$$\sum_{x=2}^{\infty} x(x-1)\alpha^{x-2}=\frac{2}{(1-\alpha)^{3}} \tag{26}$$

لاحظ إنه يمكن الحصول على (25) بتفاضل طرفي (24) ، والحصول على (26) بتفاضل طرفي (25) .

نظريـــة (6): إذا كان المتغير العشوائي X يتوزع وفق التوزيـع الهندسـي بمعلمـة p ، أي أن $X \sim Ge(p)$ فإن القيمة المتوقعة والتباين لهذا التوزيع هما على الترتيب :

$$\mu_X = E(X) = \frac{1}{p}$$

$$\sigma_X^2 = V(X) = \frac{q}{p^2}$$

البرهان :

باستخدام العلاقتين (25) و(26) يمكننا إيجاد المتوسط والتباين وذلك على النحو التالي :

$$E(X) = \sum_{x=1}^{\infty} x p q^{x-1} = p \sum_{x=1}^{\infty} x q^{x-1}$$

$$= p\left(\frac{1}{1-q}\right)^2 = \frac{p}{p^2} = \frac{1}{p}$$

وعليه فإن

$$\mu_X = E(X) = \frac{1}{p} \tag{27}$$

وبالمثل

$$E(X^2) = E[X(X-1)] + E(X)$$

حيث

$$E[X(X-1)] = \sum_{x=2}^{\infty} x(x-1) p q^{x-1} = p \sum_{x=2}^{\infty} x(x-1) q^{x-1}$$

$$= pq \sum_{x=2}^{\infty} x(x-1) q^{x-2} = pq\left[2\left(\frac{1}{1-q}\right)^3\right]$$

$$= \frac{2pq}{p^3} = \frac{2q}{p^2}$$

وبالتالي فإن

$$E(X^2) = \frac{2q}{p^2} + \frac{1}{p}$$

ولكن

$$V(X)=E(X^2)-(E(X))^2$$
$$=\frac{2q}{p^2}+\frac{1}{p}-\frac{1}{p^2}=\frac{q}{p^2}$$

إذن

$$\sigma_x^2=V(X)=\frac{q}{p^2} \quad (28)$$

مثال (12) : بفرض إنه ألقى مكعب (زهرة) نرد حتى ظهور الرقم " 1 " فإذا كان المتغير العشوائي X يمثل عدد المحاولات المطلوبة حتى ظهور الرقم " 1 " لأول مرة . فاوجد :

أ - دالة كتلة الاحتمال للمتغير العشوائي X .

ب - احتمال الحصول على الرقم " 1 " في المحاولة الثالثة .

جـ - القيمة المتوقعة والتباين للمتغير العشوائي X .

الحل :

أ - دالة احتمال المتغير العشوائي X تكون على الصيغة الأتية :

$$P(X=x)=\begin{cases}\left(\frac{5}{6}\right)^{x-1}\left(\frac{1}{6}\right) & , x=1,2,3,\ldots \\ 0 & \text{خلاف ذلك ,}\end{cases}$$

ب - احتمال الحصول على الرقم " 1 " في المحاولة الثالثة هو :

$$P(X=3)=\left(\frac{5}{6}\right)^{3-1}\left(\frac{1}{6}\right)=\frac{25}{216}$$

جـ - القيمة المتوقعة و التباين على الترتيب هما :

$$E(X)=\frac{1}{p}=6$$

و

$$V(X)=\frac{q}{p^2}=\frac{(5/6)}{(1/6)^2}=30$$

مثال (13) : وجدت الشركة العامة للالكترونيات أن 30 ٪ من المتقدمين للعمل بها يحملون شهادة البكالوريوس في الحاسوب ، و أن المتقدمين يتم اختيارهم بطريقة عشوائية و بشكل متتالي،

أ - ما احتمال أن أول شخص تتم مقابلته و يحمل شهادة البكالوريس فى الحاسوب هو الشخص الخامس .

ب - إذا كانت كل مقابله تكلف 15 دينار أوجد القيمة المتوقعة والتباين لاجمالى تكاليف المقابلة حتى يتم الحصول على العدد المطلوب من الأشخاص للوظائف الشاغرة .

الحل :

بفرض أن المتغير العشوائى X يرمز لعدد المحاولات المطلوب للحصول على أول شخص تتم مقابلته ويحمل شهادة البكالوريس فى الحاسوب . وعليه فإن المتغير العشوائى X يتوزع وفق التوزيع الهندسى بدالة كتلة إحتمال لها الصيغة الآتية :

$$P(X=x)=\begin{cases}(0.3)\,(0.7)^{x-1} & ,x=1,2,3,....\\ 0 & \text{خلاف ذلك},\end{cases}$$

أ -
$$P(X=5)=(0.3)(0.7)^4=0.072$$

ب - بفرض أن C ترمز لإجمالى التكلفه وعليه فإن :

$$C=15X$$

و منها نجد أن :

$$E(C)=15E(X)=15\left(\frac{1}{9}\right)=15\left(\frac{1}{0.3}\right)=50$$

وان :

$$V(C)=(15)^2\,V(X)=(15)^2\left(\frac{q}{p^2}\right)$$

$$=(15)^2\left(\frac{0.7}{(0.3)^2}\right)=1750$$

5 - 2 - 8 توزيع ذي الحدين السالب The Negative Binomial Distribution

لنفرض وجود متوالية لا نهائية من التجارب المستقلة حيث أن نتيجة كل تجربة إما نجاح أو فشل ، واحتمال النجاح في كل تجربة يساوي p $(0< p < 1)$ واحتمال الفشل بها يساوي $q = 1 - p$. فمثلاً عند إلقاء عملة بشكل متتال حتى الحصول على ثلاثة صور ، أو مثلاً صندوق يحتوى على n من العناصر المعيبة و m من العناصر الصالحة وسحبت منه العناصر بشكل متتال ومع الاعادة حتى الحصول على k من العناصر الصالحة ، أو مثلاً صندوق يحتوي

على m كرة بيضاء و n كرة حمراء وسحبت الكرات منه بشكل متتال ومع الاعادة حتى الحصول على اربع كرات بيضاء . . . الخ . إذن إن وجدت مثل تلك المتوالية فهي تكون متوالية لا نهائية من محاولات بيرنولي بمعلمة p ، فإذا كان المتغير العشوائي X يمثل عدد المحاولات المطلوبة للحصول على r من حالات النجاح فإن المتغير العشوائي X يتوزع وفق توزيع ذي الحدين السالب الذي دالة احتماله تأخذ الصيغة الآتية :

$$p_X(x) = P(X = x) = \begin{cases} \binom{x-1}{r-1} p^r q^{x-r} & , \; x = r, r+1, r+2, \ldots \\ 0 & , \; \text{خلاف ذلك} \end{cases} \qquad (29)$$

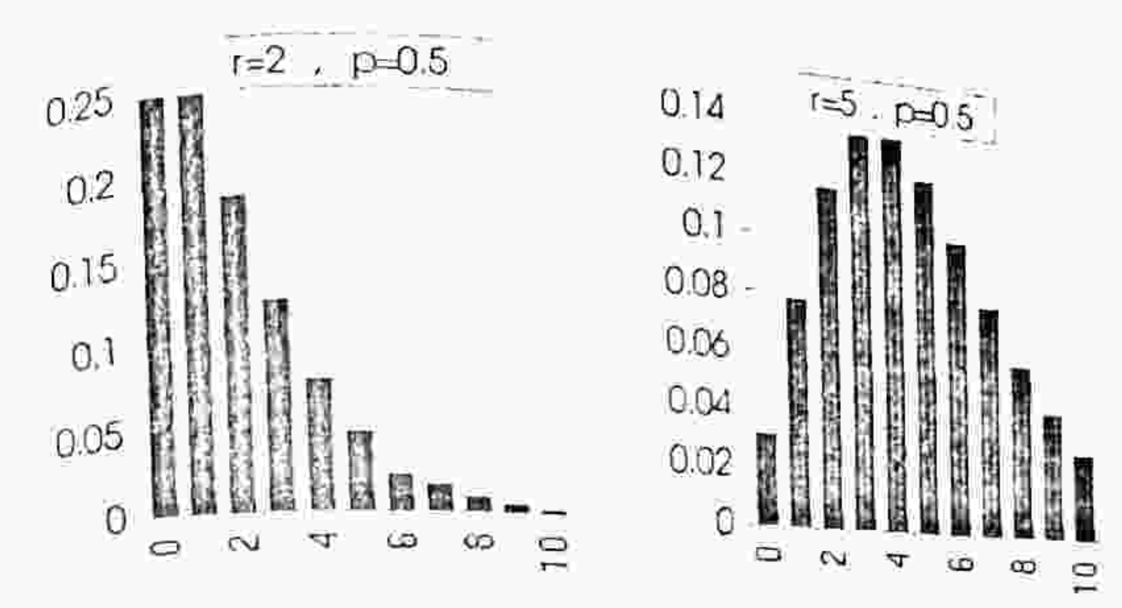
ويرمز له بالرمز $X \sim NB(r, p)$. إن التوزيع الاحتمالي للمتغير العشوائي X يعتمد على معلمتين هما r ، p . وإذا قارنا الشروط الضرورية لمتغير يتبع توزيع ذي الحدين مع الشروط الضرورية لمتغير يتبع توزيع ذي الحدين السالب نلحظ تطابق الشروط التالية :

(1) إن التجارب مستقلة .

(2) كل تجربة لها حالتان ممكنتان فقط .

(3) احتمال النجاح (p) ثابت من محاولة إلى أخرى .

ولكن الفروق الجوهرية بين التوزيعين هي أن عدد المحاولات (n) في حالة توزيع ذي الحدين ثابت ، وإن عدد حالات النجاح المشاهدة في هذه المحاولات متغير عشوائي . ولكن في حالة توزيع ذي الحدين السالب نجد أن عدد حالات النجاح (r) ثابت ولكن عدد المحاولات المطلوبة للحصول على r من حالات النجاح فهي متغير عشوائي . والاشكال الاتية تبين التمثيل البياني لدالة كتلة الاحتمال .



شكل (6) : دالة كتلة احتمال توزيع ذي الحدين السالب

عند قيم مختلفة للمعلمتين r و p

نظريـــة (7) : إذا كان المتغير العشوائي X يتوزع وفق توزيع ذي الحديـن السـالب بمعلمتين r و p أي أن $X \sim NB(r,p)$ فإن :

$$m_X = E(X) = \frac{r}{p}$$

$$\sigma_X^2 = V(X) = \frac{rq}{p^2}$$

$$m_X(t) = \left[\frac{pe^t}{1-qe^t}\right]^r$$

البرهان :

الحظ أولاً أن :

$$\binom{-r}{x} = (-1)^x \binom{r+x-1}{x} \quad , r > 0 \ , \ x \geq 0$$

وان

$$\sum_{x=0}^{\infty} \binom{-r}{x} (-q)^x = \frac{1}{(1-q)^r}$$

وعليه من تعريف القيمة المتوقعة نجد أن :

$$E(X)=\sum_{x=r}^{\infty} x\binom{x-1}{r-1}p^{r}q^{x-r}=r\,p^{r}\sum_{x=r}^{\infty}\binom{x}{r}q^{x-r}$$

$$=r\,p^{r}\sum_{y=0}^{\infty}\binom{y+r}{r}q^{y} \qquad , \; y=x-r$$

$$=r\,p^{r}\sum_{y=0}^{\infty}(-1)^{y}\binom{-r-1}{y}q^{y}$$

$$=r\,p^{r}\sum_{y=0}^{\infty}\binom{-r-1}{y}(-q)^{y}$$

$$=r\,p^{r}\cdot\frac{1}{(1-q)^{r+1}}=\frac{r}{p} \qquad (30)$$

وحيث أن :

$$E\left[X(X+1)\right]=p^{r}\sum_{x=r}^{\infty}x(x+1)\binom{x-1}{r-1}q^{x-r}$$

$$=r(r+1)\,p^{r}\sum_{y=0}^{\infty}\binom{y+r+1}{y}q^{y}$$

$$=r(r+1)\,p^{r}\sum_{y=0}^{\infty}\binom{-r-2}{y}(-q)^{y}=\frac{r(r+1)}{p^{2}}$$

وعليه فإن :

$$V(X)=E(X^{2})-\left[E(X)\right]^{2}=E\left[X(X+1)\right]-E(X)-\left[E(X)\right]^{2}$$

$$=\frac{r(r+1)}{p^{2}}-\frac{r}{p}-\left(\frac{r}{p}\right)^{2}=\frac{rq}{p^{2}} \qquad (31)$$

وأن

$$m_{X}(t)=E\left(e^{tX}\right)=\sum_{x=r}^{\infty}e^{tx}\binom{x-1}{r-1}p^{r}q^{x-r}=p^{r}\sum_{x=r}^{\infty}\binom{x-1}{x-r}\left(e^{t}\right)^{x}q^{x-r}$$

$$=p^{r}\sum_{y=0}^{\infty}\binom{y+r-1}{y}\left(e^{t}\right)^{y+r}q^{y}=p^{r}\left(e^{t}\right)^{r}\sum_{y=0}^{\infty}\binom{y+r-1}{y}\left(q\,e^{t}\right)^{y}$$

$$= \frac{(pe^t)^r}{(1-qe^t)^r} = \left[\frac{pe^t}{1-qe^t}\right]^r \qquad (32)$$

مثال (14) : بفرض أنه تم إلقاء مكعب (زهرة) نرد متزن بشكل متتال ، فما احتمال الحصول على الرقم " 6 " للمرة الثالثة في المحاولة السابعة ؟ وما هي القيمة المتوقعة والتباين؟

الحل :

إذا كان المتغير العشوائي X يمثل عدد المحاولات المطلوبة ، فإن $x = 7$ ، $p = 1/6$ ، $r = 3$

وعليه فإن الاحتمال المطلوب هو :

$$P(X=7) = \binom{6}{2}\left(\frac{1}{6}\right)^3\left(\frac{5}{6}\right)^4 = \frac{6!}{2!4!}\left(\frac{1}{6}\right)^3\left(\frac{5}{6}\right)^4$$

$$= 15\left(\frac{1}{6}\right)^3\left(\frac{5}{6}\right)^4 = 0.0335$$

وإن

$$E(X) = \frac{r}{p} = \frac{3}{1/6} = 18$$

و

$$V(X) = \frac{rq}{p^2} = \frac{3(5/6)}{(1/6)^2} = 90$$

مثال (15) : إذا علمت أن $\frac{1}{3}$ من المتبرعين بالدم يحملون الفصيلة O^+ أوجد :

أ - احتمال أن اول متبرع وفصيلة دمه O^+ يكون الشخص الرابع .

ب - احتمال أن ثاني متبرع وفصيلة دمه O^+ يكون الشخص الرابع .

الحل :

بفرض أن المتغير العشوائي X يمثل عدد المتبرعين للحصول على اول متبرع فصيلة دمه O^+ أو للحصول على ثاني متبرع وفصيلة دمه O^+ .

أ - حيث أن $x = 4$ ، $r = 1$ ، $p = \frac{1}{3}$ وعليه فإن :

$$P(X=4) = \binom{4-1}{1-1}\left(\frac{1}{3}\right)\left(\frac{2}{3}\right)^3 = 0.09877$$

ب - حيث أن $x = 4$ ، $r = 2$ ، $p = \frac{1}{3}$ وعليه فإن :

$$P(X=4)=\binom{4-1}{2-1}\left(\frac{1}{3}\right)^2\left(\frac{2}{3}\right)^2=0.14815$$

وينتهي هذا البند بملحوظة تساعد فى حساب الاحتمالات الخاصة بتوزيع ذي الحدين السالب (أو التوزيع الهندسي) ، وذلك من خلال النظر للعلاقة بين توزيع ذي الحدين السالب وتوزيع ذي الحدين وهى كالاتي :

إذا كان $X \sim B(n,p)$ و $Y \sim NB(r,p)$ فإن :

$$P(Y \le n) = P(X \ge r)$$

وإن :

$$P(Y > n) = P(X < r)$$

أي أنه إذا كان عدد المحاولات المطلوبة للحصول على حالة النجاح r أقل من أو يساوي n ، فإن عدد حالات النجاح في n من المحاولات يجب أن يكون على الأقل r . وبالتالي يمكن استخدام جداول ذي الحدين في حساب الاحتمالات الخاصة بذي الحدين السالب .

2-5-9 توزيع بواسون The Poisson Distribution

إن هذا التوزيع يكون نموذجاً احتمالياً لكثير من الظواهر العشوائية النادرة الوقوع ، فهو يستخدم في المسائل التى تتعلق بحدوث الظواهر في فترات زمنيه محدده حيث قد تكون الفترة الزمنية ثانية أو دقيقة أو ساعة أو يوماً أو اسبوعاً أو شهراً . . . الخ . كما يستخدم في المسائل التى تتعلق بحدوث الظواهر في مناطق محددة ، حيث المنطقة المحددة قد تكون صفحة من كتاب أو متراً مربعاً من مساحة . . . الخ . ومن الأمثلة على ذلك : عدد المكالمات الهاتفية التى تتلقاها بدالة كلية العلوم خلال فترة زمنية محددة ، عدد حوادث السيارات التى تحدث في طريق معين خلال يوم من ايام الاسبوع ، عدد الاهداف التى تسجل خلال مباراة في كرة السلة، عدد مرات حدوث البرق إثناء عاصفة رعدية ، عدد الأشخاص الذين يدخلون مكتب البريد كل ساعة ، عدد الاخطاء المطبعية في كتاب يحتوي على العديد من الصفحات ، عدد البكتيريا في الخلايا ، عدد الجسيمات التى تنبعث من مادة مشعة خلال فترة زمنية معينة . . . الخ .
أى أن هذا التوزيع يستخدم في وصف سلوك الاحداث النادرة بمعنى الاحداث التى تكون فيها فرصة نجاح الحدث صغيرة جداً .

يقال أن للمتغير العشوائي X توزيع بواسون بمعلمة λ ($\lambda > 0$) إذا كانت له دالة الكتلة الاحتمالية التالية :

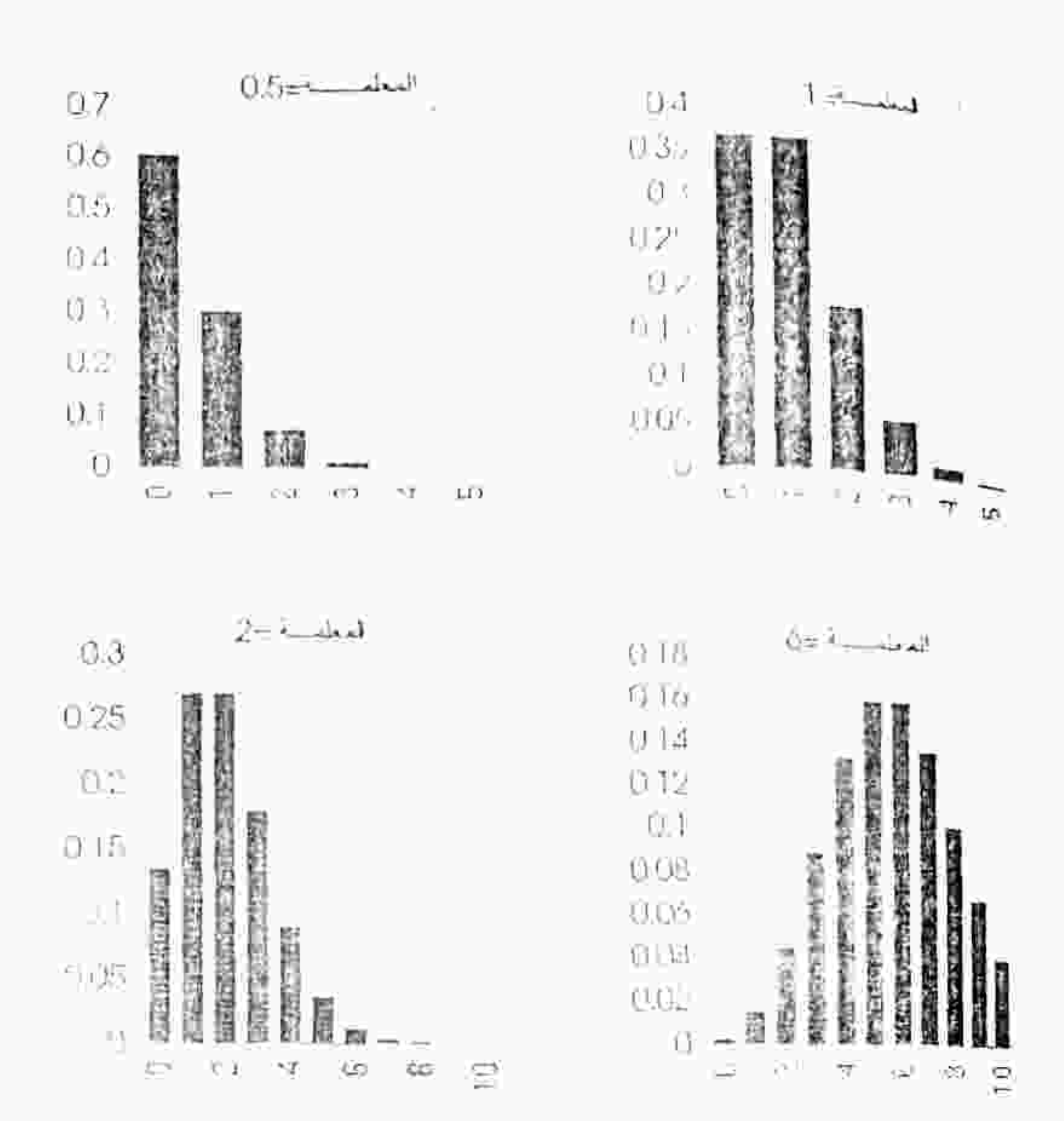
$$P(X=x)=\begin{cases} \dfrac{e^{-\lambda}\lambda^{x}}{x!} & , x=0,1,2,\ldots \\ 0 & , \text{خلاف ذلك} \end{cases} \qquad (33)$$

ويرمز لذلك بالرمز $X \sim P(\lambda)$. ومن الواضح أن $P(X=x)\geq 0$ ويمكن إثبات $\sum_{x=0}^{\infty} P(X=x)=1$ باستخدام المعلومة الرياضية التالية : $\sum_{x=0}^{\infty}\frac{a^{x}}{x!}=e^{a}$ حيث a عدد حقيقي

وعليه فإن :

$$\sum_{x=0}^{\infty} P(X=x)=\sum_{x=0}^{\infty}\frac{e^{-\lambda}\lambda^{x}}{x!}=e^{-\lambda}\sum_{x=0}^{\infty}\frac{\lambda^{x}}{x!}=e^{-\lambda}e^{\lambda}=1$$

والشكل (7) يوضح التمثيل البياني لدالة كتلة الأحتمال $P(X=x)$. ويتضح من ذلك أن د التوزيع ملتو إلى جهة اليمين عندما تكون λ صغيرة ، وهذا يتمشى مع طبيعة المتغير العشوائي X وذلك لأن احتمال فشله كبير واحتمال نجاحه صغير .



شكل (7) : دالة كتلة الاحتمال لتوزيع بواسون عند قيم مختلفة للمعل

نظريــة (8) : إذا كان المتغير العشوائي X يتوزع وفق توزيع بو

$$= E(X) = \lambda$$

$$V(X) = \lambda$$

البرهان :

من تعريف القيمة المتوقعة ... أن ...

$$E(X)=\sum_{x=0}^{\infty} x P(X=x)=\sum_{x=0}^{\infty} x \frac{e^{-\lambda} \lambda^{x}}{x!}=e^{-\lambda} \sum_{x=1}^{\infty} \frac{\lambda \lambda^{x-1}}{(x-1)!}$$

$$=\lambda e^{-\lambda} \sum_{y=0}^{\infty} \frac{\lambda^{y}}{y!} \qquad , y=x-1$$

$$=\lambda e^{-\lambda} e^{\lambda}=\lambda$$

إذن

$$\mu_{x}=E(X)=\lambda \qquad (34)$$

وإن

$$E[X(X-1)]=\sum_{x=0}^{\infty} x(x-1) \frac{e^{-\lambda} \lambda^{x}}{x!}=e^{-\lambda} \sum_{x=2}^{\infty} \frac{x(x-1) \lambda^{2} \lambda^{x-2}}{x(x-1)(x-2)!}$$

$$=e^{-\lambda} \lambda^{2} \sum_{x=2}^{\infty} \frac{\lambda^{x-2}}{(x-2)!}=e^{-\lambda} \lambda^{2} \sum_{y=0}^{\infty} \frac{\lambda^{y}}{y!} \qquad , y=x-2$$

$$=e^{-\lambda} \lambda^{2} e^{\lambda}=\lambda^{2}$$

وعليه فإن

$$E\left(X^{2}\right)=E[X(X-1)]+E(X)=\lambda^{2}+\lambda$$

ومنها نجد أن

$$\sigma_{X}^{2}=V(X)=E\left(X^{2}\right)-[E(X)]^{2}=\lambda^{2}+\lambda-\lambda^{2}=\lambda \qquad (35)$$

ومن تعريف الدالة المولدة للعزوم يتضح أن :

$$m_{x}(t)=E\left(e^{tX}\right)=\sum_{x=0}^{\infty} e^{tx} \frac{e^{-\lambda} \lambda^{x}}{x!}=e^{-\lambda} \sum_{x=0}^{\infty} \frac{\left(\lambda e^{t}\right)^{x}}{x!}=e^{-\lambda} e^{\lambda e^{t}}$$

$$=e^{\lambda\left(e^{t}-1\right)}=\exp \left(\lambda\left(e^{t}-1\right)\right) \qquad (36)$$

مثال (16) : إذا علمت أن عدد حوادث السيارات التي تحدث في الأسبوع بمدينة معينة تتبع توزيع بواسون بمتوسط يساوي 0.7 فما احتمال وقوع ثلاثة حوادث على الأقل خلال الأسبوع؟

الحل :

إذا كان المتغير العشوائي X يمثل عدد حوادث السيارات التي تحدث في الأسبوع فإن X يتوزع وفق توزيع بواسون بمعلمة $\lambda=0.7$ وعليه فإن :

$$P(X \geq 3) = \sum_{x=3}^{\infty} \frac{e^{-0.7}(0.7)^x}{x!}$$

$$= 1 - P(X < 3) = 1 - P(X \leq 2)$$

$$= 1 - \sum_{x=0}^{2} \frac{e^{-0.7}(0.7)^x}{x!}$$

$$= 1 - [0.4966 + 0.3476 + 0.1217] = 1 - 0.9659 = 0.0341$$

لاحظ أنه بالإمكان الحصول على هذا الاحتمال باستخدام جدول احتمالات توزيع بواسون في اخر هذا الكتاب (جدول رقم (3)) ، فعندما $\lambda = 0.7$ نجد أن :

$$P(X \geq 3) = 1 - P(X < 3)$$

$$= 1 - [P(X = 0) + P(X = 1) + P(X = 2)]$$

$$= 1 - [0.4966 + 0.3476 + 0.1217] = 0.0341$$

مثال (17) : إذا علمت بأن معدل العمليات الجراحية بأحد الأقسام بمركز طرابلس الطبي 3 في اليوم فاحسب احتمال وقوع الأحداث التالية :

أ - عدم اجراء اى عملية جراحية في يوم معين .

ب - اجراء عملية جراحية واحدة على الاقل في يوم معين

ج - اجراء عمليتين جراحيتين على الاكثر في يوم معين .

د - اجراء من 4 إلى 6 عمليات جراحية خلال يومين .

الحل :

إذا كان المتغير العشوائي X يمثل عدد العمليات الجراحية في اليوم الواحد فإن هذا المتغير يتوزع وفق توزيع بواسون بمعلمة $\lambda = 3$ وعليه فإن :

$$P(X = x) = \begin{cases} \frac{e^{-3}3^x}{x!} & , x = 0, 1, 2, \dots \\ 0 & \text{خلاف ذلك .} \end{cases}$$

وبالتالي بالامكان الحصول على الاحتمالات المطلوبة باستخدام دالة كتلة الاحتمال لتوزيع بواسون او باستخدام جدول توزيع بواسون رقم (3) .

أ - احتمال عدم اجراء اى عملية جراحية في يوم معين : $P(X = 0)$

ومن الجدول رقم (3) عندما $\lambda = 3$ و $x = 0$ نجد أن : $P(X = 0) = 0.0498$

ب - احتمال اجراء عملية جراحية واحدة على الاقل في يوم معين : $P(X \geq 1)$

وعليه فإن :
$$P(X \geq 1) = 1 - P(X = 0) = 1 - 0..0498 = 0.9502$$

جـ - احتمال اجراء عمليتين جراحيتين على الاكثر في يوم معين : $P(X \leq 2)$

وبالتالي فإن :
$$P(X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2)$$
$$= 0.0498 + 0.1494 + 0.2240 = 0.4232$$

د - حيث أن معدل العمليات الجراحية في اليوم يساوي 3 فإن معدلها خلال يومين يساوي 6

اي أن : $\lambda = 2 \times 3 = 6$

احتمال اجراء من 4 إلى 6 عمليات جراحية خلال يومين : $P(4 \leq X \leq 6)$

وباستخدام جدول رقم (3) عندما $\lambda = 6$ و $x = 4,5,6$ نجد أن :
$$P(4 \leq X \leq 6) = P(X = 4) + P(X = 5) + P(X = 6)$$
$$= 0.1339 + 0.1606 + 0.1606 = 0.4551$$

5 - 2 - 10 تقريب توزيع ذي الحدين بتوزيع بواسون :

The Poisson Approximation to The Binomial Distribution

يمكن اشتقاق توزيع بواسون من توزيع ذي الحدين وذلك كحالة تقريبية تحت شروط معينة وهو ما سنتناوله في هذا البند . فتوزيع بواسون يعتبر حالة خاصة من توزيع ذي الحدين وذلك عندما يكون احتمال النجاح صغير جداً $(p \to 0)$ ، وعدد المحاولات (n) كبيراً $(n \to \infty)$ بحيث تبقى $np = \lambda$ ثابتة . فإذا كان المتغير العشوائي X يتوزع وفق توزيع ذي الحدين وتتوفر فيه الشروط التالية :

(I) عدد المحاولات (n) كبيراً ، غالباً $n \geq 50$.

(II) احتمال النجاح (p) صغير جداً ، في الغالب $p \leq 0.10$.

(III) حاصل ضرب np أقل من 5 ، أي أن $np \leq 5$.

فإن هذا المتغير يتوزع توزيعاً قريباً من توزيع بواسون . أي أن :
$$P(X = x) = \binom{n}{x} p^x q^{n-x} \cong \frac{e^{-\lambda} \lambda^x}{x!} \quad , \lambda = np$$

إن لهذه الخاصية أهمية خاصة من الناحية التطبيقية ، حيث إنها تسمح باستخدام توزيع بواسون بديلاً لتوزيع ذي الحدين عند حساب احتمال حدوث حدث معين ، وذلك بمجرد التأكد من أن n كبيرة ($n \geq 50$) واحتمال النجاح صغير جداً ($p \leq 0.10$) .

مما سبق يمكن تلخيص خواص توزيع بواسون في النقاط التالية :

1 - توزيع بواسون توزيع احتمالي منفصل .

2 - المتوسط الحسابي والتباين متساويين وكل منهما يساوي λ .

3 - يتحدد هذا التوزيع بمعلومية المعلمة λ ، فبمعرفتها نستطيع تحديد القيم التي من الممكن أن يأخذها المتغير العشوائي X واحتمال كل منها .

4 - بالإمكان إثبات أن المعامل العزمي للالتواء يساوي $\frac{1}{(\lambda)^{1/2}}$. وبذلك فإن المعامل العزمي للالتواء في توزيع بواسون لا يمكن أن يكون سالباً أو مساوياً صفر ، وهذا دليل على أن التوزيع غير متماثل وملتوي التواء موجب (إلى اليمين) .

5 - كذلك يمكن إثبات أن المعامل العزمي للتفرطح يساوي $1/\lambda + 3$. وهذا يدل على أن توزيع بواسون يميل دائماً نحو التذبذب خاصة لقيم λ الصغيرة .

مثال (18) : تلتزم الشركة العامة للإلكترونات بإصلاح الأجهزة خلال مدة ستة أشهر من تاريخ بيعها ، فإذا علمت أن احتمال حدوث عطل بأي جهاز خلال هذه المدة هو 0.02 فما احتمال أن من بين 300 جهاز تم بيعها سوف تلتزم الشركة بإصلاح على الأقل أربعة منها ؟

الحل :

حيث إن $n = 300$ كبيرة و $p = 0.02$ صغيرة ، وعليه نرى من الأفضل استخدام توزيع بواسون في حساب هذا الاحتمال لأن استخدام توزيع ذي الحدين سيكون معقداً . فإذا كان المتغير العشوائي X يمثل عدد الأجهزة التي حدث بها عطل فإن :

$$\lambda = n\,p = 300\,(0.02) = 6$$

وبذلك يكون الاحتمال المطلوب هو :

$$P(X \ge 4) = \sum_{x=4}^{\infty} \frac{e^{-6} 6^x}{x!}$$

$$= 1 - P(X < 4) = 1 - P(X \le 3)$$

$$= 1 - \sum_{x=0}^{3} \frac{e^{-6} 6^x}{x!}$$

$$= 1 - [0.0025 + 0.0149 + 0.0446 + 0.0894]$$

$$= 1 - 0.1502 = 0.8498$$

مثال (19) : إذا كان 10 % من إنتاج مصنع معين لسلعة ما معيباً ، وتم اختيار عينة تتكون من 50 وحدة من إنتاج ذلك المصنع ، فما احتمال وجود وحدتين معيبتين ؟

الحل :

بفرض أن المتغير العشوائي X يمثل عدد الوحدات المعيبة ، وحيث أن $p = 0.10$ و $n = 50$ فإن $\lambda = np = 5$.

إذن

$$P(X = 2) = \frac{e^{-5} 5^2}{2!} = 0.0842$$

الحظ انه إذا استخدمنا توزيع ذي الحدين نجد أن :

$$P(X = 2) = \binom{50}{2} (0.10)^2 (0.90)^{48} = 0.0779$$

وهذا الاختلاف في الناتج يرجع لاستخدام التقريب .

5 - 3 توزيعات متصلة (مستمرة) Continuous Distributions

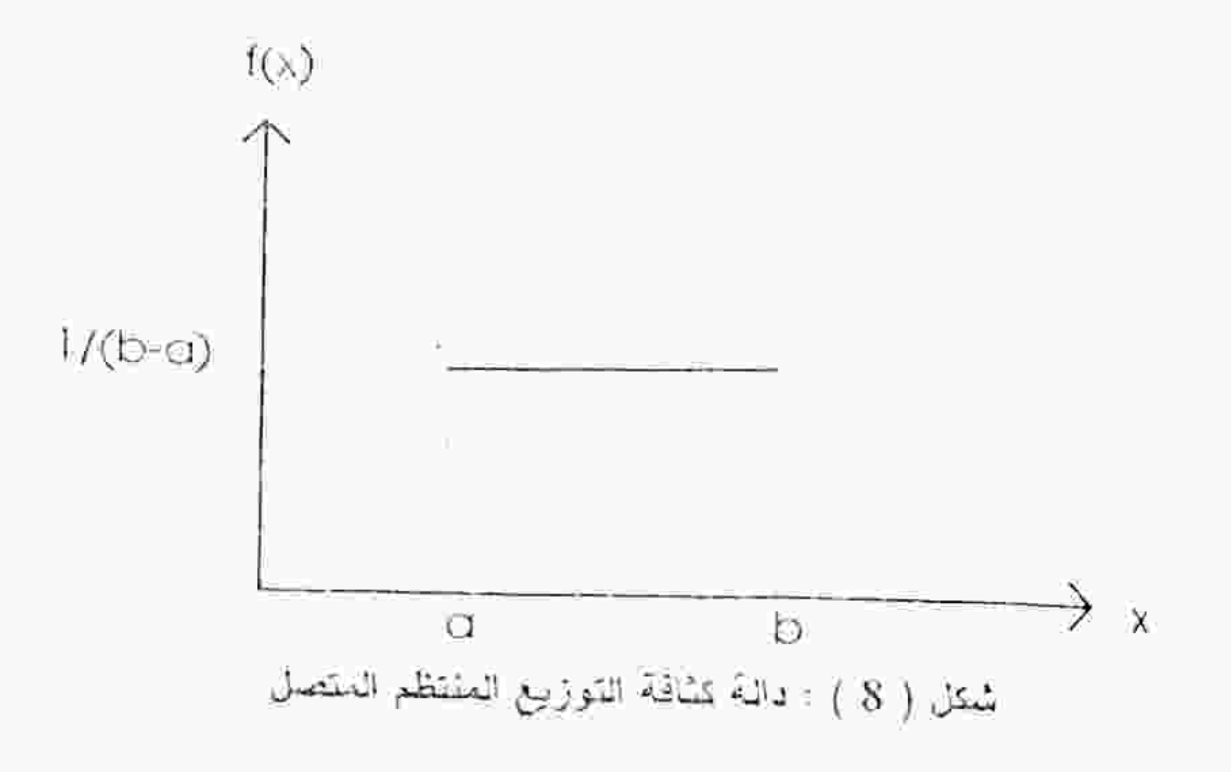
لقد تعرضنا في ما سبق للتوزيعات المنفصلة وفي هذا البند سوف نعرض عدة توزيعات متصلة مع اشتقاق كل من المتوسط والتباين والدالة المولدة للعزوم (متى كان ذلك ممكناً) .

5 - 3 - 1 التوزيع المنتظم المتصل Continuous Uniform Distribution

إذا كان X متغيراً عشوائياً معرف على الفترة $[a, b]$ وبدالة كثافة احتمال معرفة كالآتي :

$$f_X(x)=\begin{cases}\dfrac{1}{b-a} & , a\le x\le b \\ 0 & \text{خلاف ذلك} ,\end{cases} \quad (37)$$

فإن المتغير العشوائي X يتوزع وفق التوزيع المنتظم بمعلمتين a و b . ويرمز لذلك بالرمز $X \sim U_{[a,b]}$ ، والشكل (8) يبين التمثيل البياني لدالة كثافة احتمال المتغير العشوائي X .



شكل (8) : دالة كثافة التوزيع المنتظم المتصل

ولما

$$\int_a^b f_X(x)\,dx=\int_a^b \frac{1}{b-a}\,dx=1$$

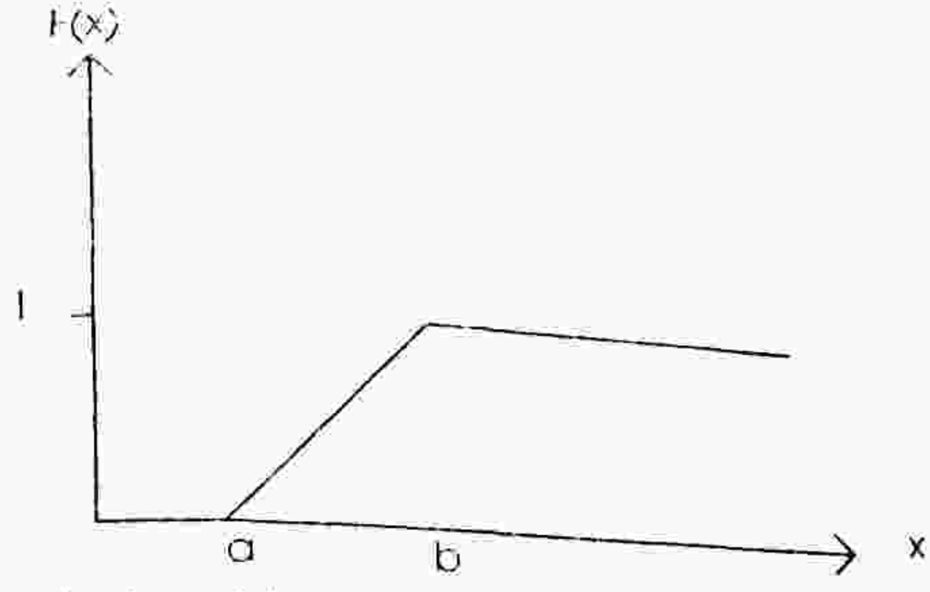
وكذلك دالة التوزيع التراكمي للمتغير X تكون كما يلي :

$$F_X(x)=P(X\le x)=\int_a^x \frac{1}{b-a}\,dt=\frac{x-a}{b-a}$$

وعليه فإن

$$F_X(x)=\begin{cases}0 & , x<a \\ \dfrac{x-a}{b-a} & , a\le x\le b \\ 1 & , x\ge b\end{cases} \quad (38)$$

والشكل (9) يبين التمثيل البياني لدالة التوزيع التراكمي .



شكل (9) : دالة التوزيع التراكمي للمتغير العشوائي المنتظم

نظرية (9): إذا كان المتغير العشوائي X يتوزع وفق التوزيع المنتظم المتصل بمعلمتين a و b
أي أن $X \sim U[a,b]$ فإن :

$$\mu_X = E(X) = \frac{a+b}{2}$$

$$\sigma_X^2 = V(X) = \frac{(b-a)^2}{12}$$

$$m_X(t) = \frac{e^{bt} - e^{at}}{t(b-a)}$$

البرهان :

من تعريف القيمة المتوقعة للمتغير العشوائي المتصل نجد أن :

$$E(X) = \int_a^b x\, f_X(x)\, dx = \int_a^b x \frac{1}{b-a}\, dx$$

$$= \frac{1}{b-a}\left[\frac{x^2}{2}\right]_a^b = \frac{b^2-a^2}{2(b-a)}$$

$$= \frac{a+b}{2} \qquad (39)$$

وأن

$$E(X^2)=\int_a^b \frac{x^2}{b-a}\,dx=\frac{1}{b-a}\int_a^b x^2\,dx$$

$$=\frac{1}{b-a}\cdot\frac{b^3-a^3}{3}=\frac{b^3-a^3}{3(b-a)}=\frac{a^2+ab+b^2}{3}$$

وعليه فإن :

$$V(X)=E(X^2)-[E(X)]^2=\frac{a^2+ab+b^2}{3}-\frac{(a+b)^2}{4}$$

$$=\frac{(b-a)^2}{12} \qquad (40)$$

ومن تعريف الدالة المولدة للعزوم يتضح أن :

$$m_X(t)=E(e^{tX})=\int_a^b \frac{e^{tx}}{b-a}\,dx=\frac{1}{b-a}\left[\frac{e^{bt}-e^{at}}{t}\right]$$

$$=\frac{e^{bt}-e^{at}}{t(b-a)} \quad , t>0 \qquad (41)$$

إن لهذا التوزيع تطبيقات مهمة في الإحصاءات اللامعلمية وفي التقليد (المحاكاة) Simulation، حيث يستخدم في توليد (generate) عينات عشوائية من التوزيعات المتصلة .

مثال (20) : إذا كان المتغير العشوائي X يتوزع وفق التوزيع المنتظم في الفترة [10 , 0] فاحسب الاحتمالات التالية :

1) $P(X<3)$ 2) $P(X>6)$ 3) $P(3<X<8)$

الحل :

حيث أن : $a=0$ و $b=10$ في الصيغة العامة للتوزيع المنتظم ، وعليه فإن :

$$f_X(x)=\begin{cases}\frac{1}{10} & , 0\le x\le 10\\ 0 & \text{، خلاف ذلك}\end{cases}$$

وبالتالي فإن :

1 - $$P(X<3)=\int_0^3 \frac{1}{10}\,dx=\frac{3}{10}$$

2 - $$P(X>6)=\int_6^{10} \frac{1}{10}\,dx=\frac{4}{10}$$

$$P(3<X<8)=\int_{3}^{8}\frac{1}{10}\,dx=\frac{1}{2}$$

- 3

مثال (21) : إذا كانت الحافلات تصل إلى محطة معينة كل 15 دقيقة ابتداء من الساعة السابعة صباحاً ، بمعنى تصل على تمام الساعة 7:00 ، 7:15 ، 7:30 ، 7:45 وهكذا . ووصل أحد الركاب إلى تلك المحطة في زمن يتبع التوزيع المنتظم بين الساعة السابعة والسابعة ونصف [7:00 ، 7:30] ، أوجد احتمال أن هذا الراكب سوف ينتظر الحافلة

(I) أقل من 5 دقائق .   (II) أكثر من 10 دقائق .

الحل :

بجعل المتغير العشوائي X يمثل زمن وصول الراكب بعد الساعة السابعة صباحاً إلى تلك المحطة، نجد أن X يتوزع وفق التوزيع المنتظم خلال الفترة (0 , 30) .

وبذلك فإن هذا الراكب يجب أن ينتظر أقل من 5 دقائق إذا وفقط إذا وصل بين الساعة 7:10 و 7:15 أو بين الساعة 7:25 و 7:30 ، ويكون الاحتمال المطلوب في (I) هو :

$$P(10<X<15)+P(25<X<30)=\int_{10}^{15}\frac{1}{30}\,dx+\int_{25}^{30}\frac{1}{30}\,dx=\frac{1}{3}$$

وبالمثل يجب أن ينتظر أكثر من 10 دقائق إذا وصل بين الساعة 7:00 و 7:05 أو بين الساعة 7:15 و 7:20 ، وبذلك يكون الاحتمال المطلوب في (II) هو :

$$P(0<X<5)+P(15<X<20)=\int_{0}^{5}\frac{1}{30}\,dx+\int_{15}^{20}\frac{1}{30}\,dx=\frac{1}{3}$$

5 . 3 . 2 التوزيع الأسي The Exponential Distribution

غالباً ما يستخدم هذا التوزيع في المسائل العملية الخاصة بزمن انتظار وقوع حدث معين . فمثلاً الزمن الذي يبقى فيه جهاز إلكتروني أو آلة بمصنع صالحة للعمل قبل أن يحدث بها عطب يعطلها عن العمل ، أو مثلاً الفترة الزمنية التي ينتظرها زبون بمقهى أو بمصرف قبل أن تقدم إليه أي خدمة ، أو الفترة الزمنية ما بين وصول زبونين لمصرف أو لمطعم أو إلى أي مكان آخر تقدم فيه خدمات . فإذا كان وقوع الحوادث يقع وفقاً لنظام بواسون ، فإن زمن الانتظار لوقوع حدث معين والفترة الزمنية ما بين حدثين متتاليين سيكون لهما توزيع أسي . وهذه الحقيقة مفيدة جداً في استخدام التوزيع الأسي في كثير من المسائل العملية ، ويعرف التوزيع الأسي كما يلي :

يقال للمتغير العشوائي X توزيع أسي بمعلمة λ (λ > 0) إذا كانت دالـة كثافة احتماله لها الصيغة الآتية :

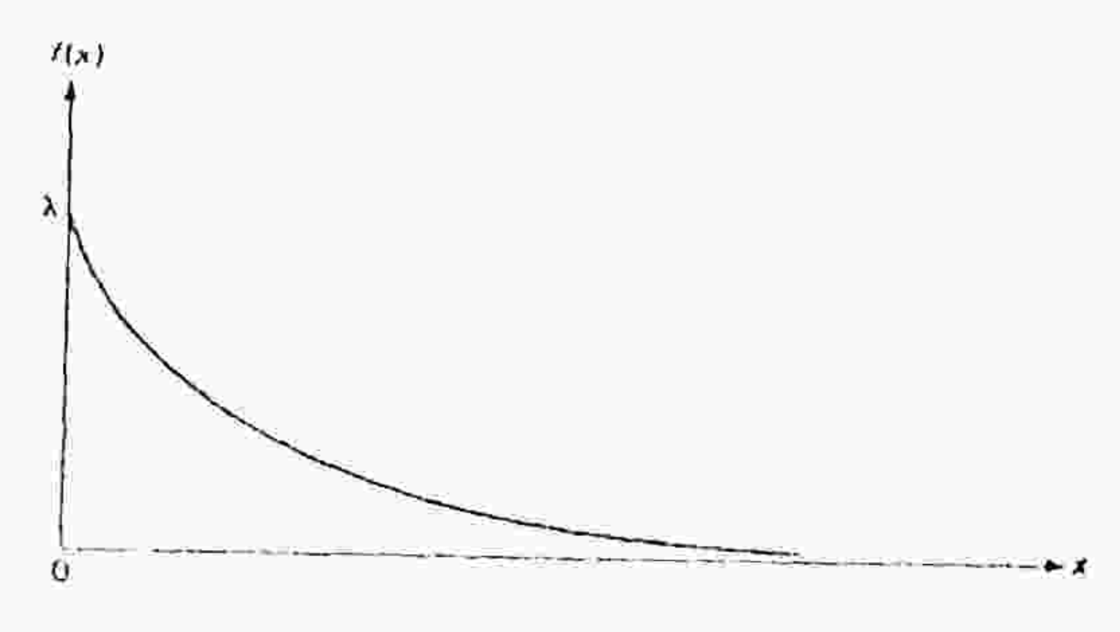
$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & , \ x \geq 0 \\ 0 & , \ x < 0 \end{cases} \qquad (42)$$

ويرمز له بالرمز X ~ Exp (λ) . ودالة التوزيع التراكمي لها الصيغة التالية :

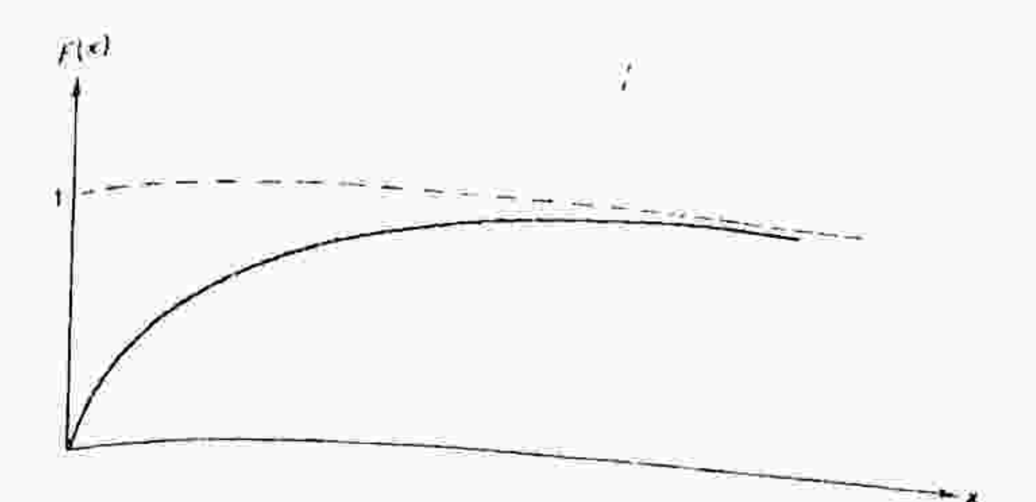
$$F_X(x) = \begin{cases} 0 & , \ x < 0 \\ 1 - e^{-\lambda x} & , \ x \geq 0 \end{cases} \qquad (43)$$

والشكل (10) يمثل دالة الكثافة بيانياً وذلك عندما $\lambda = 1$

والشكل (11) يمثل دالة التوزيع التراكمية لهذا المتغير العشوائي .



شكل (10) : دالة الكثافة الاحتمالية للمتغير العشوائي الأسي



شكل (11) : دالة التوزيع التراكمية للمتغير العشوائي الأسي

نظريــة (10) : إذا كان المتغير العشوائي X يتوزع وفق التوزيع الأسي بمعلمة λ فإن :

$$\mu_x = E(X) = \frac{1}{\lambda}$$

$$\sigma_x^2 = V(X) = \frac{1}{\lambda^2}$$

$$m_x(t) = \frac{\lambda}{\lambda - t} \quad , t < \lambda$$

البرهان :

إنه من الأسهل حساب الدالة المولدة للعزوم ومنها يتم إيجاد المتوسط والتباين ، وعليه فإن

$$m_x(t) = E\left(e^{tx}\right) = \int_0^\infty e^{tx} \lambda e^{-\lambda x} dx$$

$$= \lambda \int_0^\infty e^{-x(\lambda - t)} dx = \frac{\lambda}{\lambda - t} \quad , t < \lambda \qquad (44)$$

إذن

ومنها

$$m_x'(t) = \frac{\partial m_x(t)}{\partial t} = \frac{\lambda}{(\lambda - t)^2}$$

(45)

$$m''_X(t)=\frac{2\lambda}{(\lambda-t)^3} \quad \text{ومنها} \quad \begin{aligned} m'_X(0)&=E(X)=\frac{1}{\lambda} \\ m''_X(0)&=E(X^2)=\frac{2}{\lambda^2} \end{aligned}$$

ويكون فإن

$$V(X)=E(X^2)-[E(X)]^2=\frac{2}{\lambda^2}-\frac{1}{\lambda^2}=\frac{1}{\lambda^2} \qquad (46)$$

بصفة عامة فإن :

$$m_X^{(r)}(t)=\frac{r!\,\lambda}{(\lambda-t)^{r}} \quad \Rightarrow \quad m_X^{(r)}(0)=\frac{r!}{\lambda^r} \quad , r=1,2,\ \ldots$$

أي أن

$$E(X^r)=\frac{r!}{\lambda^r} \quad , r=1,2,\ \ldots \qquad (47)$$

إن لهذا التوزيع خاصية تميزه عن بقية التوزيعات المتصلة الأخرى وهي إذا كان s و t عددين حيث $s>0$ و $t>0$ فإن :

$$P(X>t+s/X>t)=\frac{P(X>t+s)}{P(X>t)}=\frac{e^{-\lambda(t+s)}}{e^{-\lambda t}}$$

$$=e^{-\lambda s}=P(X>s) \qquad (48)$$

إن ما تتضمنه الصيغة (48) ، هو أنه إذا لم يحدث الحدث في الزمن t فإن احتمال عدم حدوثه في الزمن القادم s هو $e^{-\lambda s}$. وهو نفس احتمال حدوث الحدث في الزمن s ابتداء من الزمن 0 هذه الخاصية تسمى فقدان الذاكرة .

مثال (22) : نفرض أن الزمن الذي يقضيه الزبون بأحد المصارف يتبع التوزيع الأسي بمتوسط 4 دقائق ، فإذا تم اختيار زبون بشكل عشوائي من ذلك المصرف فأوجد :

أ - احتمال انتظاره أكثر من 10 دقائق .

ب - احتمال أن الزبون سينتظر اكثر من 10 دقائق علما بأنه انتظر اكثر من 5 دقائق .

الحل :

نفترض أن المتغير العشوائي X يمثل زمن الانتظار للزبون بالمصرف ، وان $\mu_x = 5$ وعليه

فإن : $\lambda = \frac{1}{5}$.

أ - احتمال انتظاره أكثر من 10 دقائق هو :

$$P(X>10) = e^{-10\lambda} = e^{-10(1/5)} = e^{-2} = 0.1353$$

ب - احتمال أن الزبون سينتظر اكثر من 10 دقائق علما بأنه انتظر اكثر من 5 دقائق هو :

$$P(X>10 / X>5) = P(X>5) = e^{-5(1/5)} = e^{-1} = 0.3679$$

مثال (23) : إذا كان الزمن الذي تستغرقه مكالمة هاتفية بالدقائق بأحد الإدارات يتوزع وفق التوزيع الأسي بمتوسط 5 دقائق ، وتم اختيار إحدى هذه المكالمات بطريقة عشوائية فأوجد احتمال وقوع الأحداث الآتية :

أ - أن تستغرق هذه المكالمة أكثر من دقيقتين .

ب - أن تستغرق هذه المكالمة أقل من دقيقتين .

جـ - أن تستغرق هذه المكالمة من 5 إلى 10 دقائق .

الحل :

أ - احتمال أن تستغرق هذه المكالمة أكثر من دقيقتين :

$$P(X>2) = \int_{2}^{\infty} \frac{1}{5} e^{-\frac{1}{5}x}\, dx = -e^{-\frac{1}{5}x} \Big|_{2}^{\infty} = e^{-\frac{2}{5}} = e^{-0.4} = 0.6703$$

ب - احتمال أن تستغرق هذه المكالمة أقل من دقيقتين :

$$P(X<2) = 1 - P(X \geq 2) = 1 - e^{-\frac{2}{5}} = 1 - 0.6703 = 0.3297$$

جـ - احتمال أن تستغرق هذه المكالمة من 5 إلى 10 دقائق :

$$P(5 \leq X \leq 10) = \int_{5}^{10} \frac{1}{5} e^{-\frac{1}{5}x} dx = -e^{-\frac{1}{5}x} \Big|_{5}^{10}$$

$$= -e^{-2} + e^{-1} = -0.1353 + 0.3679 = 0.2326$$

3.3.5 التوزيع الطبيعي The Normal Distribution

إن التوزيع الطبيعي الذي سنناقشه في هذا البند يعد من أهم التوزيعات الاحتمالية في الإحصاء ويرجع ذلك للأسباب الآتية : أولاً إن كثيراً من الظواهر التي تظهر في التجارب العملية تتوزع توزيعاً طبيعياً ، ثانياً في بعض الاحيان قد لا يكون المتغير العشوائي موزعاً توزيعاً طبيعياً ولكن يمكن أن يحول إلى متغير عشوائي يتوزع توزيعاً طبيعياً وذلك تحت شروط معينة ، ثالثاً قد توجد توزيعات معقدة وبالتالي يمكن أن يستخدم التوزيع الطبيعي تقريباً لها ، وأخيراً إن غالبية التوزيعات الاحتمالية منفصلة كانت أم متصلة يتقارب توزيعها من التوزيع الطبيعي وذلك من خلال نظرية النهاية المركزية Central Limit Theorem . وإن هذا التوزيع يدخل في كافة المجالات الصناعية والزراعية والاقتصادية وغيرها من المجالات الأخرى . وفيما يلي تعريف هذا التوزيع .

إذا كان X متغيراً عشوائياً يتوزع وفق التوزيع الطبيعي فإن دالة كثافة احتماله لها الصيغة الآتية:

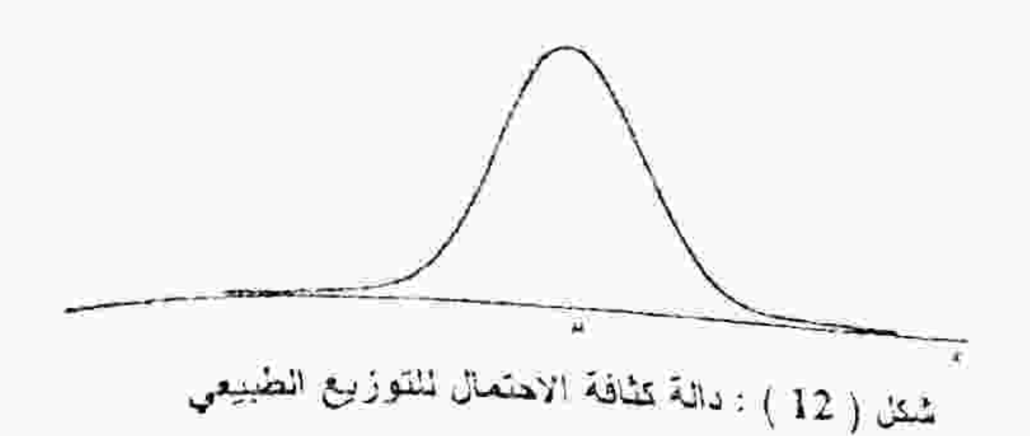
$$f_X(x)=\frac{1}{\sqrt{2\pi\sigma^2}}\,e^{-\frac{1}{2\sigma^2}(x-\mu)^2} \qquad ,-\infty<x<\infty \qquad (49)$$

حيث μ $(-\infty<\mu<\infty)$ و σ^2 $(\sigma^2>0)$ يمثلان معلمتي التوزيع وهما المتوسط والتباين على التوالي (كما سنرى فيما بعد) .

ويرمز للتوزيع الطبيعي عادة بالرمز $N(\mu\,,\sigma^2)$ ، ونقول بأن $X\sim N(\mu\,,\sigma^2)$.

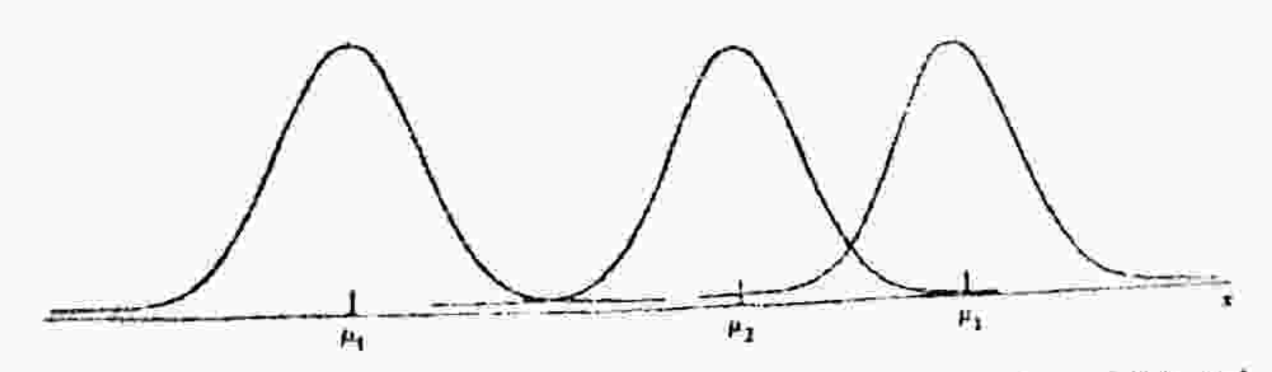
وبالتمعن في دالة الكثافة الاحتمالية وتمثيلها البياني يتضح أن :

1- منحنى التوزيع الطبيعي ناقوسي الشكل ومتماثل وتقع قمته فوق المتوسط ويمتد طرفاه إلى ما لانهاية من الجانبين دون أن يلامسا المحور الأفقي . وبالتحديد إن دالة الكثافة متماثلة حول $x=\mu$ أي أن $f(\mu-x)=f(\mu+x)$ وهذا يعني أن المتوسط يساوي الوسيط ويساوي المنوال وأن نقطتي الانقلاب في منحنى الدالة هما $x_1=\mu-\sigma$ و $x_2=\mu+\sigma$ وانهما يقعان على بعد ثابت على يمين ويسار المنوال ، كما في شكل (12) .



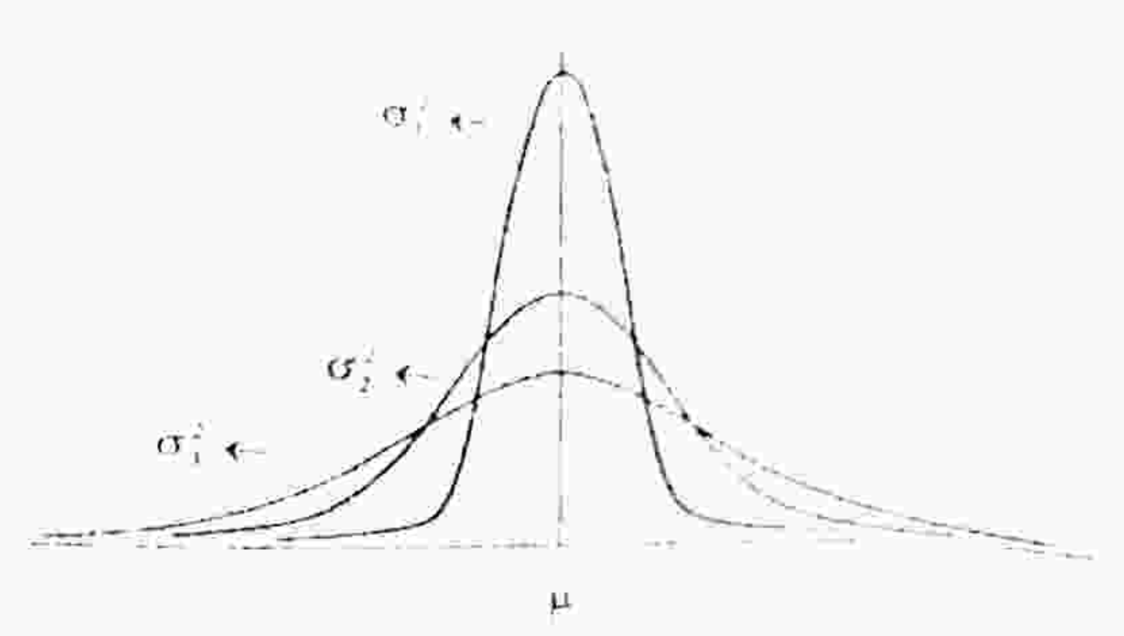
شكل (12) : دالة كثافة الاحتمال للتوزيع الطبيعي

2 - يتحدد التوزيع بمعلومية μ و σ^2 ، ويختلف التوزيع إذا اختلفت μ أو σ^2 أو كلاهما ، حيث أن قيمة μ تحدد موقع التوزيع الطبيعي على الخط الحقيقي ، فكلما زادت هذه القيمة تغير موقع المنحنى في الاتجاه الأيمن والعكس صحيح . بينما σ^2 تبين مقدار تشتت وتفرطح منحنى الدالة . فكلما كانت σ^2 صغيرة كلما كان المنحنى مدبباً ، وكلما كانت كبيرة كلما كان المنحنى متفرطح بمعنى آخر كلما قل التباين ارتفعت قمة المنحنى وزاد تقارب الطرفين والعكس صحيح وذلك كما هو موضح في الأشكال (13) و (14) و (15) .



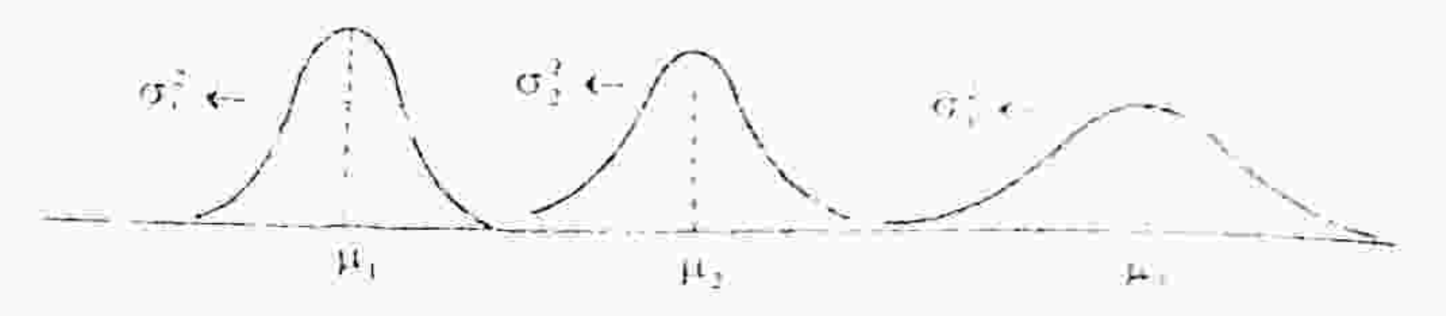
شكل (13) : ثلاث توزيعات طبيعية متساوية التباينات ومختلفة المتوسطات

(دالة كثافة الاحتمال لثلاث توزيعات طبيعية عندما $\mu_1 < \mu_2 < \mu_3$ بينما $\sigma_1^2 = \sigma_2^2 = \sigma_3^2$)



شكل (14) : ثلاث توزيعات طبيعية متساوية المتوسطات ومختلفة التباين .

(دالة كثافة احتمال ثلاث توزيعات طبيعية عندما $\mu_1 = \mu_2 = \mu_3 = \mu$ و $\sigma_1^2 < \sigma_2^2 < \sigma_3^2$)



شكل (15) : ثلاث توزيعات طبيعية مختلفة المتوسطات والتباينات .

(دالة كثافة الاحتمال لثلاث توزيعات طبيعية عندما $\mu_1 < \mu_2 < \mu_3$ و $\sigma_1^2 < \sigma_2^2 < \sigma_3^2$)

3- المساحة تحت منحنى التوزيع الطبيعي تساوي واحد صحيح ، أي أن :

$$\int_{-\infty}^{\infty} f_X(x)\, dx = 1$$

وكما هو موضح في شكل (16) أدناه :



شكل (16) . المساحة المظللة = 1

ويمكن إثبات ذلك على النحو التالي : بوضع $y = \frac{x-\mu}{\sigma}$ نجد أن

$$\int_{-\infty}^{\infty} f_X(x)\,dx = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}}\, e^{-\frac{y^2}{2}}\, dy \qquad (50)$$

وبوضع

$$I = \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}}\, dy \qquad (51)$$

يجب إثبات أن $I = \sqrt{2\pi}$. إذن من المعادلة (51) نجد أن :

$$I^2 = I.I = \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}}\, dy \int_{-\infty}^{\infty} e^{-\frac{z^2}{2}}\, dz$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-\frac{1}{2}(y^2+z^2)}\, dy\, dz$$

سوف نغير الآن المتغيرات بهذا التكامل من y و z إلى الإحداثيات القطبية r و θ (Polar Coordinates) وذلك بوضع $y = r\cos\theta$ و $z = r\sin\theta$ وعليه فإن $r > 0$ ، $y^2 + z^2 = r^2$ و $0 < \theta < 2\pi$ وبالتالي فإن :

$$I^2 = \int_0^{2\pi}\int_0^{\infty} e^{-\frac{r^2}{2}}\, r\, dr\, d\theta = 2\pi$$

إذن $I = \sqrt{2\pi}$ وعليه فإن

$$\int_{-\infty}^{\infty} f_X(x)\, dx = 1$$

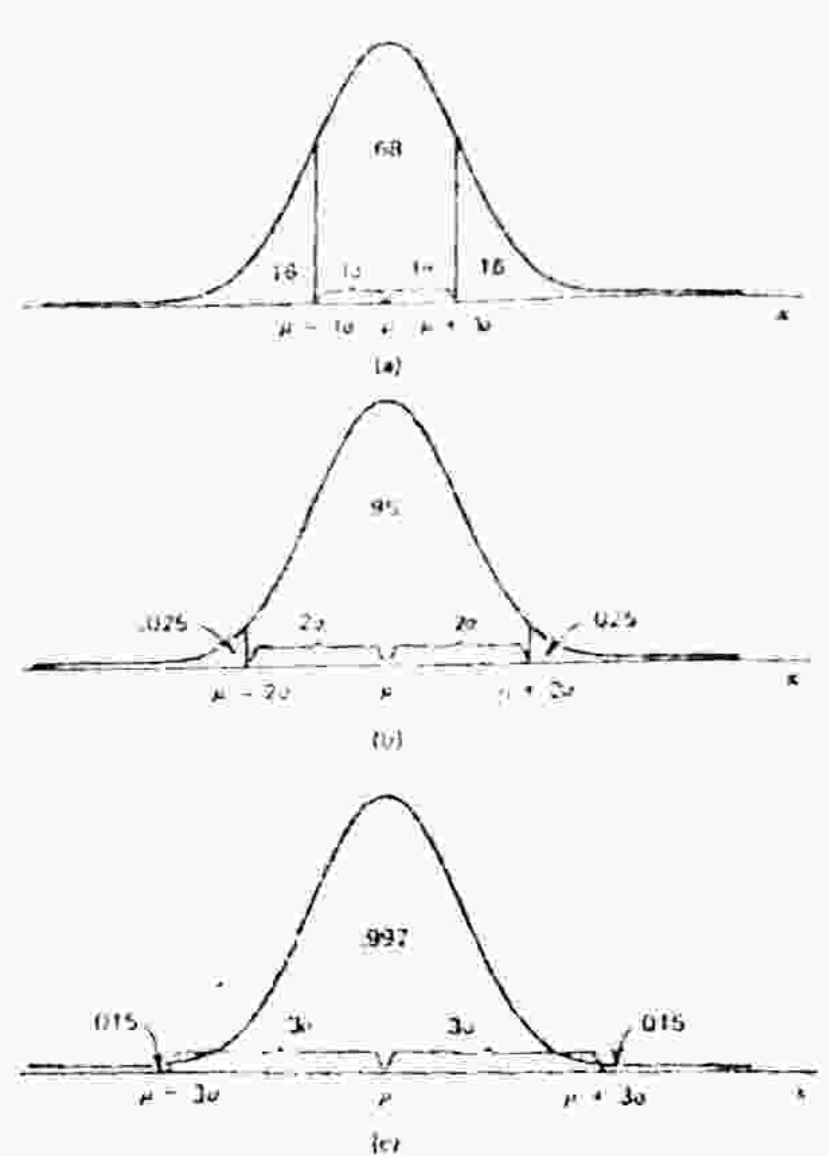
...نسب معينة من المساحة الواقعة ضمن أي عدد من الانحراف المعياري فمثلاً :

المساحة الواقعة ضمن انحراف معياري واحد عن المتوسط تساوي ٦٨٪ من المساحة الكلية.

المساحة الواقعة ضمن انحرافين معياريين عن المتوسط تساوي ٩٥٪ من المساحة الكلية .

المساحة الواقعة ضمن ثلاثة انحرافات معيارية عن المتوسط تساوي ٩٩٪ من المساحة الكلية

يوضح الحالات الثلاثة السابقة بيانياً كما في شكل (17) التالي :



شكل (17) : النسب المئوية من المساحة الواقعة ضمن عدد معين من الانحراف المعياري

- معامل العزمي للالتواء يساوي صفر لجميع المنحنيات الطبيعية ، وذلك لكونها متماثلة حول محور تماثلها تحت قمة المنحنى .
- معامل العزمي للتفرطح يساوي 3 لجميع المنحنيات الطبيعية متوسطة التفرطح .

نظرية (11) : إذا كان المتغير العشوائي X يتوزع وفق التوزيع الطبيعي بمعلمتين μ و σ^2 فإن

$$\mu_X = E(X) = \mu$$

$$\sigma_X^2 = V(X) = \sigma^2$$

$$m_X(t) = e^{\mu t + \frac{1}{2}\sigma^2 t^2}$$

البرهان :

لقد ذكرنا عند تعريفنا لدالة كثافة الاحتمال بأن المعلمتين μ و σ^2 تمثلان متوسط وتباين هذا التوزيع ، ولتبرير استخدام هذه المصطلحات ، يجب إثبات أن μ فعلاً تمثل المتوسط و σ^2 تمثل التباين ، وعليه سوف نوجد كل منهما على النحو التالي :

$$E(X) = \int_{-\infty}^{\infty} x f_X(x)\, dx = \int_{-\infty}^{\infty} x \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}\, dx$$

$$= \int_{-\infty}^{\infty} (\sigma y + \mu) \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}\, dy \qquad , y = \frac{x-\mu}{\sigma}$$

$$= \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y e^{-\frac{y^2}{2}}\, dy + \mu \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}\, dy$$

ولكن الحد الثاني من هذه المتساوية عبارة عن توزيع طبيعي بمتوسط يساوي صفر وتباين يساوي واحد ، وحيث أن :

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}\, dy = 1 \qquad \Rightarrow \qquad \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}}\, dy = \sqrt{2\pi}$$

ولكن $g(y) = y e^{-\frac{y^2}{2}}$ دالة فردية أي أن $g(-y) = -g(y)$ ، وعليه فإن :

$$\int_{-\infty}^{\infty} y e^{-\frac{y^2}{2}}\, dy = 0$$

علاوة على ذلك . $\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}}\, dy$ لأنها تمثل دالة توزيع طبيعي بمتوسط $\mu = 0$ وتباين $\sigma^2 = 1$ وعليه فإن :

فإن

$$E(X) = \frac{\sigma}{\sqrt{2\pi}} \times 0 + \mu \times 1 = \mu$$

$$E(X^2)=\int_{-\infty}^{\infty} x^2 \cdot f_X(x)\,dx=\frac{1}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{\infty} x^2 \quad e^{-\frac{1}{2\sigma^2}(x-\mu)^2}\,dx$$

وبوضع $y=\frac{x-\mu}{\sigma}$ نجد أن :

$$E(X^2)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}(\sigma y+\mu)^2\ e^{-\frac{y^2}{2}}$$

$$=\frac{\sigma^2}{\sqrt{2\pi}}\int_{-\infty}^{\infty} y^2\ e^{-\frac{y^2}{2}}+2\mu\sigma\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} y\ e^{-\frac{y^2}{2}}+\mu^2\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{y^2}{2}}dy$$

وحيث أن

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{y^2}{2}}\ dy=1 \quad \text{و} \quad \int_{-\infty}^{\infty} y\ e^{-\frac{y^2}{2}}\ dy=0$$

وبالتالي فإن :

$$E(X^2)=\frac{\sigma^2}{\sqrt{2\pi}}\int_{-\infty}^{\infty} y^2\ e^{-\frac{y^2}{2}}\ +\mu^2$$

وإنه بوضع $u=y$ و $dv=y\ e^{-\frac{y^2}{2}}dy$ نجد أن $v=-e^{-\frac{y^2}{2}}$ و $du=dy$

وبالتالي يكون :

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} y^2\ e^{-\frac{y^2}{2}}=\frac{1}{\sqrt{2\pi}}\left[-y\,e^{-\frac{y^2}{2}}\right]_{-\infty}^{\infty}+\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{y^2}{2}}dy=0+1=1$$

ومنها نجد أن :

$$E(X^2)=\sigma^2+\mu^2$$

عليه فإن :

$$V(X)=E(X^2)-[E(X)]^2=E(X^2)-\mu^2=(\sigma^2+\mu^2)-\mu^2=\sigma^2$$

إن الدالة المولدة للعزوم هي :

$$m_X(t)=E\left(e^{tX}\right)=\int_{-\infty}^{\infty} e^{tx}\ f_X(x)\ dx$$

$$=\int_{-\infty}^{\infty} e^{tx}\ \frac{1}{\sqrt{2\pi\sigma^2}}\ e^{-\frac{1}{2\sigma^2}(x-\mu)^2}\ dx$$

$$=\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}}\ e^{\left[tx-\frac{1}{2\sigma^2}(x-\mu)^2\right]}\ dx$$

وبإكمال المربع داخل القوس نحصل على العلاقة الآتية :

$$tx-\frac{1}{2\sigma^2}(x-\mu)^2=\mu t+\frac{\sigma^2 t^2}{2}-\frac{\left[x-\left(\mu+\sigma^2 t\right)\right]^2}{2\sigma^2}$$

وعليه فإن

$$m_X(t)=c\,e^{\mu t+\frac{1}{2}\sigma^2 t^2}$$

حيث

$$c=\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}}\ e^{-\frac{1}{2\sigma^2}\left[x-\left(\mu+\sigma^2 t\right)\right]^2}\ dx$$

وإذا استبدلنا μ بالمقدار $\mu+\sigma^2 t$ في المعادلة (49) ومن حقيقة أن $\int f_X(x)\,dx=1$ نجد أن $c=1$. وعليه فإن الدالة المولدة للعزوم التوزيع الطبيعي لها الصيغة الآتية :

$$m_X(t)=e^{\mu t+\frac{1}{2}\sigma^2 t^2} \qquad -\infty<t<\infty \qquad (52)$$

5 - 3 - 4 التوزيع الطبيعي المعياري The Standard Normal Distribution

عندما يكون متوسط التوزيع الطبيعي يساوي صفراً والتباين يساوي واحداً فإنه يسمى التوزيع الطبيعي المعياري ، وعادةً يرمز لدالة كثافة احتمال التوزيع الطبيعي المعياري بالرمز ϕ ولدالة التوزيع التراكمي بالرمز Φ . وعليه إذا كان X متغيراً عشوائياً طبيعياً معيارياً فإن :

$$\phi(x)=f_X(x)=\frac{1}{\sqrt{2\pi}}\,e^{-\frac{x^2}{2}} \quad , -\infty<x<\infty \qquad (53)$$

$$\Phi(x)=F_X(x)=\int_{-\infty}^{x}\phi(u)\,du \quad , -\infty<x<\infty \qquad (54)$$

...الة التوزيع التراكمي $\Phi(x)$ لا يمكن وضعها في صيغة سهلة ، وبالتالي فإن احتمالات التوزيع الطبيعي المعياري أو أي توزيع طبيعي آخر يتم حسابها باستخدام القيم الجدولية للدالة $\Phi(x)$ كتلك الموجودة في آخر هذا الكتاب . والنظرية الآتية توضح كيفية حساب الاحتمال لمتغير عشوائي يتبع التوزيع الطبيعي بمتوسط μ وتباين σ^2 .

نظريـة (12) : إذا كان $X \sim N(\mu, \sigma^2)$ فإن :

$$P(a < X < b) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right) \quad (55)$$

البرهـان :

$$P(a<X<b) = \int_a^b f_X(x)\,dx$$

$$= \int_a^b \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}\,dx$$

$$= \int_{\left(\frac{a-\mu}{\sigma}\right)}^{\left(\frac{b-\mu}{\sigma}\right)} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}\,dz \quad , z = \frac{x-\mu}{\sigma}$$

$$= \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$$

وحيث أن دالة التوزيع الطبيعي المعياري متماثلة حول النقطة $x = 0$ وعليه فإن :

$$P(X \leq x) = P(X \geq -x)$$

...منها يتضح أن :

$$\Phi(x) = 1 - \Phi(-x) \quad , -\infty < x < \infty \quad (56)$$

إن جدول التوزيع الطبيعي المعياري الموجود في آخر هذا الكتاب يعطي $P(0 < Z < a)$ ...حيث $Z \sim N(0,1)$ أما الاحتمالات المناظرة للقيم التالية فيتم حسابها بالتماثل ، وعليه لأي ...a يكون :

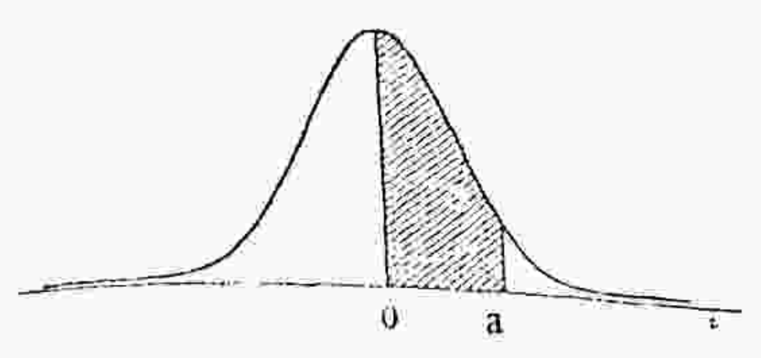
$$P(-\infty < Z < a) = 0.5 + P(0 < Z < a)$$

$$P(0 < Z < a) = P(-a < Z < 0)$$

$$P(|Z| < a) = P(-a < Z < a) = 2P(0 < Z < a)$$

$$P(|Z| > a) = 2P(Z > a) = 2[0.5 - P(0 < Z < a)]$$

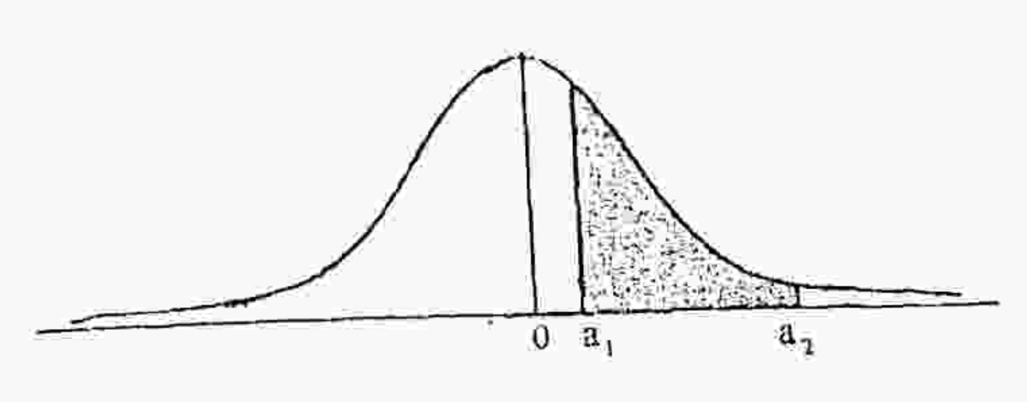
وشكل (18) يوضح المساحة تحت منحنى التوزيع الطبيعي المحصورة بين 0 و a ، $P(0 < Z < a)$.



شكل (18) : المساحة المضللة = $P(0 < Z < a)$

وبالمثل $P(a_1 < Z < a_2) = P(Z < a_2) - P(Z \leq a_1)$

وذلك كما هو مبين في شكل (19) .

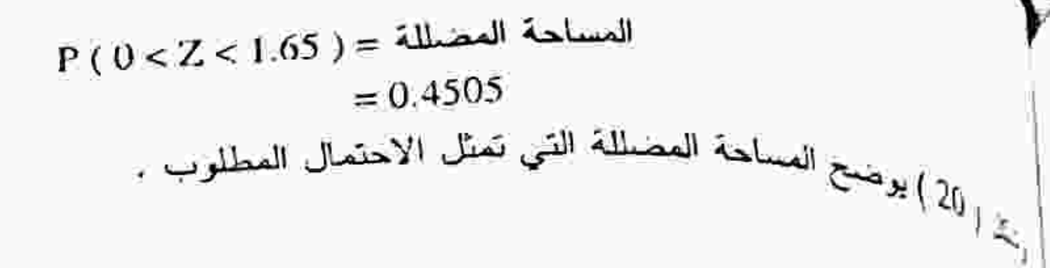


شكل (19) : المساحة المضللة = $P(a_1 < Z < a_2)$

مثال (24) : إذا كان المتغير العشوائي Z يتبع التوزيع الطبيعي المعياري، فأوجد : $P(0 < Z < 1.65)$.

الحل :

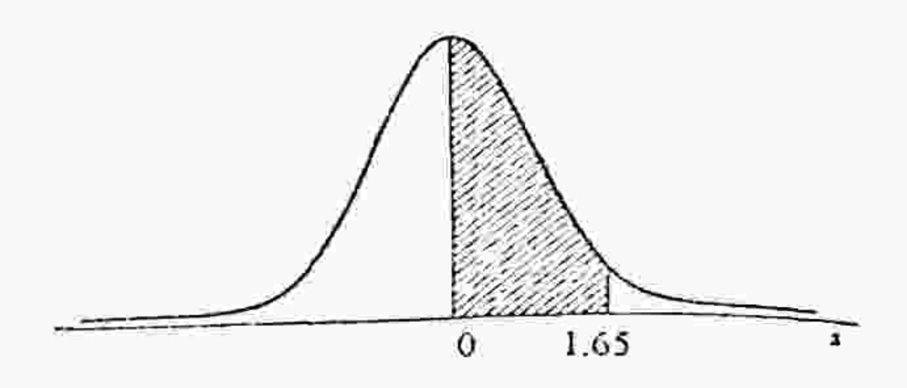
باستخدام جدول المساحات للتوزيع الطبيعي المعياري بآخر هذا الكتاب ، وذلك بالبحث في العمود الهامشي عن العدد 1.6 وفي السطر الهامشي عن العدد 0.05 نجد أن :



المساحة المضللة = $P(0 < Z < 1.65)$

$= 0.4505$

شكل (20) يوضح المساحة المضللة التي تمثل الاحتمال المطلوب .



شكل (22): المساحة المضللة = $P(0 < Z < 1.65)$

مثال (25) : إذا علمت أن درجات طلبة الدراسات العليا في امتحان القبول (GRE) تتوزع وفق التوزيع الطبيعي بمتوسط 500 درجة وانحراف معياري 100 درجة فما نسبة الطلبة الذين يحصلون على درجة أعلى من 643 درجة ؟

الحل :

إذا افترضنا أن المتغير العشوائي X يمثل درجات الطلبة في ذلك الامتحان فإن الاحتمال المطلوب هو : $P(X > 643)$. وبالتالي فإن :

$$P(X > 643) = P\left(Z > \frac{643 - 500}{100}\right) = P(Z > 1.43) = 0.5 - P(0 < Z < 1.43)$$

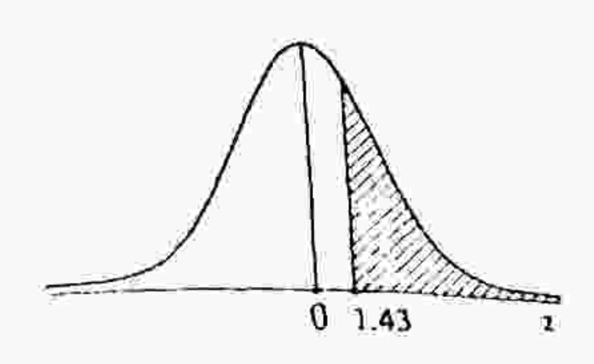
من جدول التوزيع الطبيعي رقم (4) نجد أن :

$$P(0 < Z < 1.43) = 0.4236$$

وعليه فإن :

$$P(X > 643) = 0.5 - P(0 < Z < 1.43) = 0.5 - 0.4236 = 0.0764$$

وذلك يعني أن 7.64 ٪ من الطلبة تحصلوا على درجة أعلى من 643 درجة ، بينما 92.36 ٪
منهم تحصل على درجة اقل من 643 . والشكل (21) يبين الدرجات الأصلية والمعيارية
بينما المساحة المضللة تمثل الاحتمال المطلوب .



شكل (21) : المساحة المضللة = $P(X > 643) = P(Z > 1.43)$

مثال (26):في المثال السابق ماهي الدرجة التي تحصل 5٪ من الطلبة على درجة أعلى منها؟

الحل :

بافتراض أن الدرجة التي تحصل 5٪ من الطلبة على درجة أعلى منها هي c ، فإن المطلوب إيجاد قيمة c التي تحقق الاحتمال التالي :

$$P(X > c) = 0.05$$

وعليه فإن :

$$P(X > c) = P\left(Z > \frac{c-500}{100}\right) = P(Z > z_0) = 0.05$$

حيث

(a)

$$z_0 = \frac{c-500}{100}$$

ومن جدول التوزيع الطبيعي رقم (4) نجد أن :

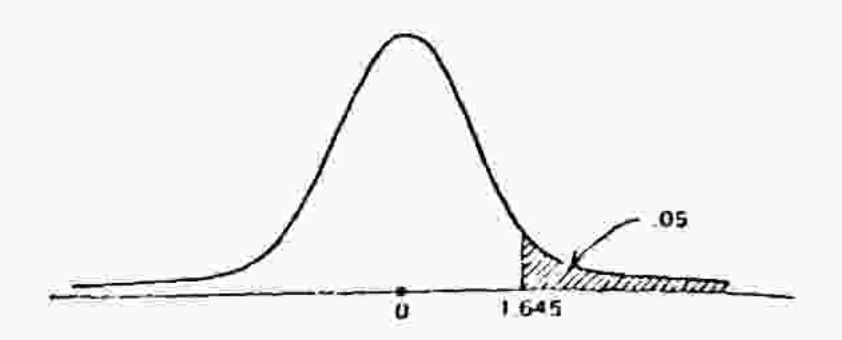
(b)

$$z_0 = 1.645$$

من (a) و (b) يتضح أن :

$$\frac{c-500}{100} = 1.645 \quad \Rightarrow \quad c = 500 + 100 \times 1.645 = 664.5$$

وبالتالي فإن الدرجة التي تحصل 5٪ من الطلبة على درجة أعلى منها هي 664.5



شكل (22) : المساحة المضللة = $P(X > 664.5) = P(Z > 1.645)$

مثال (27) : ينتج مصنع تاجوراء للنضائد والإطارات ، نضائد صغيرة متوسط عمرها الزمني يساوي 76 ساعة بانحراف معياري يساوي 10 ساعات ، فإذا علمت أن العمر الزمني لهذه النضائد يتبع التوزيع الطبيعي ، فأوجد احتمال أن يكون العمر الزمني لنضيدة ما :

أ- بين 71 و 82 ساعة . ب - اكثر من 75 ساعة . جـ - اقل من 78 ساعة .

الحل :

بفرض أن المتغير العشوائي X يمثل العمر الزمني للنضائد الصغيرة التي ينتجها مصنع تاجوراء للنضائد والإطارات فإن : $X \sim N(\mu = 76 \ , \ \sigma^2 = 100)$

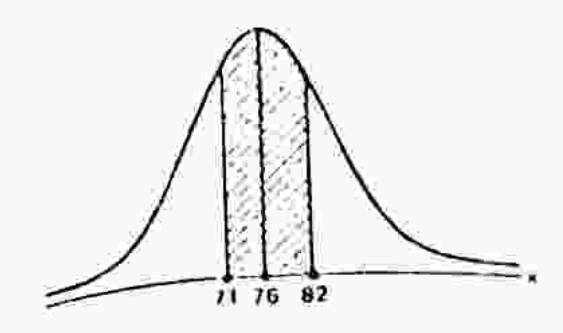
أ- بتحويل المتغير العشوائي X الذي يتوزع وفق التوزيع الطبيعي إلى متغير عشوائي يتبع التوزيع الطبيعي المعياري (Z) حتى يتسنى لنا استخدام جدول التوزيع الطبيعى المعياري ، نجد أن :

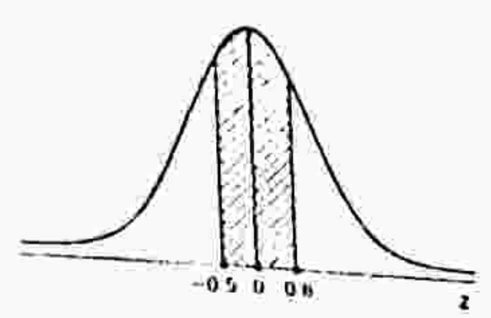
$$P(71 \leq X \leq 82) = P\left(\frac{71-\mu}{\sigma} \leq Z \leq \frac{82-\mu}{\sigma}\right) = P\left(\frac{71-76}{10} \leq Z \leq \frac{82-76}{10}\right)$$

$$= P(-0.5 \leq Z \leq 0.6) = P(-0.5 \leq Z \leq 0) + P(0 \leq Z \leq 0.6)$$

$$= 0.1915 + 0.2257 = 0.4172$$

وذلك باستخدام جدول التوزيع الطبيعي رقم (4) . والشكل التالي يوضح المساحة التي تمثل الاحتمال المطلوب باستخدام التوزيع الطبيعي الأصلي والتوزيع الطبيعي المعياري .





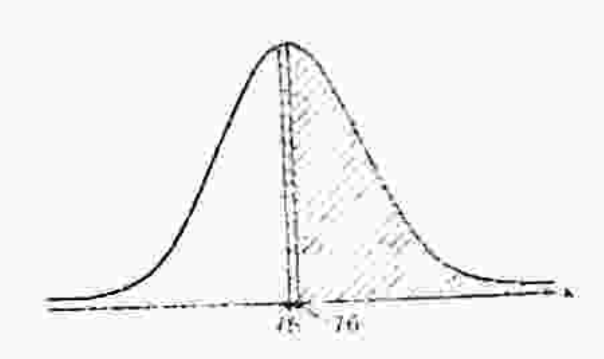
شكل (23) : المساحة المضللة = $P(71 \le X \le 82) = P(-0.6 \le Z \le 0.5)$

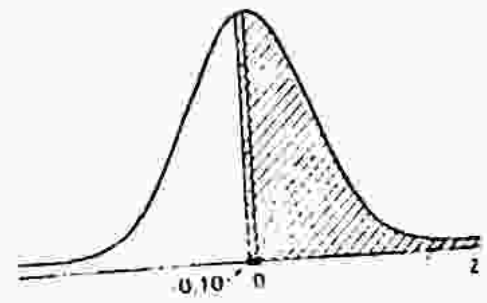
ب - الاحتمال المطلوب هو :

$$P(X > 75) = P\left(Z > \frac{75-\mu}{\sigma}\right) = P\left(Z > \frac{75-76}{10}\right) = P(Z > -0.10)$$

$$= P(-0.10 \le Z \le 0) + 0.5 = 0.0398 + 0.5 = 0.5398$$

والشكل التالي يوضح الاحتمال المطلوب في صورة مساحة .





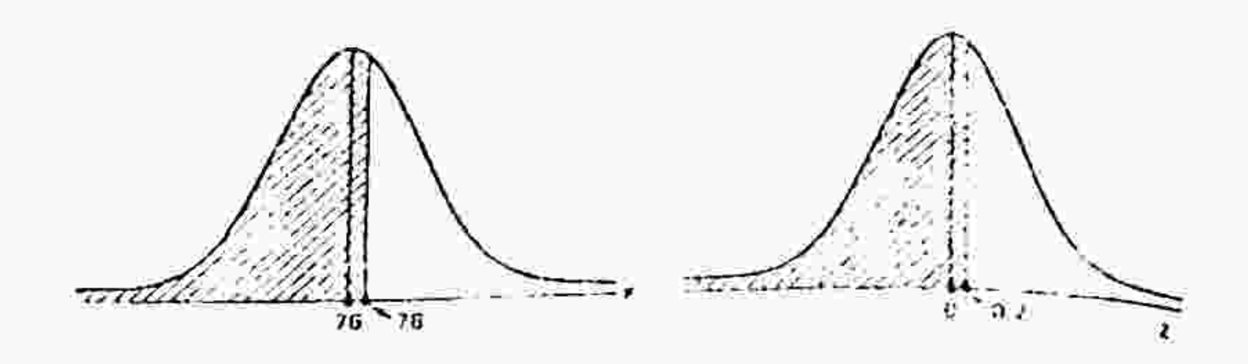
شكل (24) : المساحة المضللة = $P(X > 75) = P(Z > -0.10)$

جـ - الاحتمال المطلوب هو :

$$P(X < 78) = P\left(Z < \frac{78-\mu}{\sigma}\right) = P\left(Z < \frac{78-76}{10}\right) = P(Z < 0.2)$$

$$= 0.5 + P(0 \le Z \le 0.2) = 0.5 + 0.0793 = 0.5793$$

والشكل (25) يوضح المساحة التي تمثل الاحتمال المطلوب باستخدام التوزيع الطبيعي الأصلي والتوزيع الطبيعي المعياري .



شكل (25) : المساحة المضللة = $P(X < 78) = P(Z < 0.2)$

مثال (28) : إذا كان المتغير العشوائي X يتوزع وفق التوزيع الطبيعي بمتوسط يساوي 4 وانحراف معياري يساوي 0.012 فأوجد قيمة الاحتمال التالي : $P(3.97 \le X \le 4.03)$.

الحل :

$$P(3.97 \le X \le 4.03) = P\left(\frac{3.97-\mu}{\sigma} \le Z \le \frac{4.03-\mu}{\sigma}\right) = P\left(\frac{3.97-4}{0.012} \le Z \le \frac{4.03-4}{0.012}\right)$$

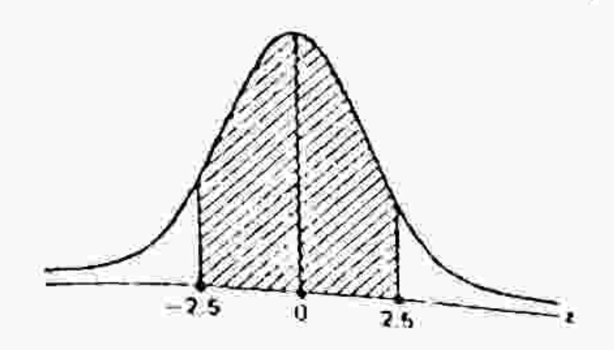
$$= P(-2.5 \le Z \le 2.5)$$

$$= P(-2.5 \le Z \le 0) + P(0 \le Z \le 2.5)$$

$$= 2P(0 \le Z \le 2.5) \quad \text{(من خاصية التماثل)}$$

$$= 2 \times 0.4938 = 0.9876$$

وذلك كما هو موضح في الشكل التالي :



شكل (26) : المساحة المضللة = $P(3.97 \leq X \leq 4.03) = P(-2.5 \leq Z \leq 2.5)$

مثال (29): إذا كانت أطوال تلاميذ وتلميذات مدرسة شهداء رأس الغزال للتعليم الأساسي تتوزع وفق التوزيع الطبيعي بمتوسط 120 سم وتباين 16 فكم عدد التلاميذ والتلميذات الذين تزيد أطوالهم عن 121 سم ، علماً بأن العدد الكلي للتلاميذ والتلميذات بهذه المدرسة يساوي 500 .

الحل :

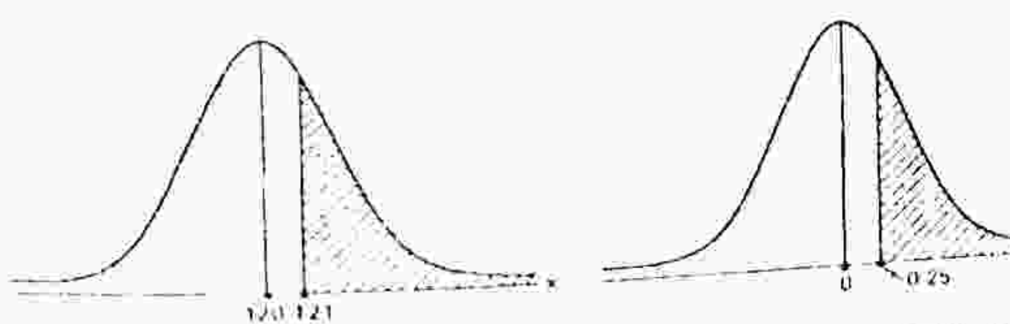
بفرض أن المتغير العشوائي X يمثل أطوال تلاميذ وتلميذات مدرسة رأس الغزال وبالتالي فإن :

$$X \sim N(\mu = 120 \ , \ \sigma^2 = 16)$$

وبذلك فإن نسبة التلاميذ والتلميذات الذين تزيد أطوالهم عن 121 سم يمثلها الاحتمال التالي :

$$P(X > 121) = P\left(Z > \frac{121-\mu}{\sigma}\right) = P\left(Z > \frac{121-120}{4}\right) = P(Z > 0.25)$$

$$= 0.5 - P(0 \leq Z \leq 0.25) = 0.5 - 0.0987 = 0.4013$$



شكل (27) : المساحة المضللة = $P(X > 121) = P(Z > 0.25)$

وبالتالي فإن العدد المتوقع للتلاميذ والتلميذات (k) الذين أطوالهم تزيد عن 121 سم يساوي تقريباً :

$$k = 0.4013 \times 500 = 201$$

مثال (30) : إذا كان وزن الخراف خلال الثلاثة الأشهر الأولى من عمرها بأحد مشاريع تربية الأغنام يتبع في تغيراته التوزيع الطبيعي بمتوسط 14 كجم وانحراف معياري 1.22 كجم ، فما هي نسبة الخراف التي يقل وزنها عن 12 كجم ؟

الحل :

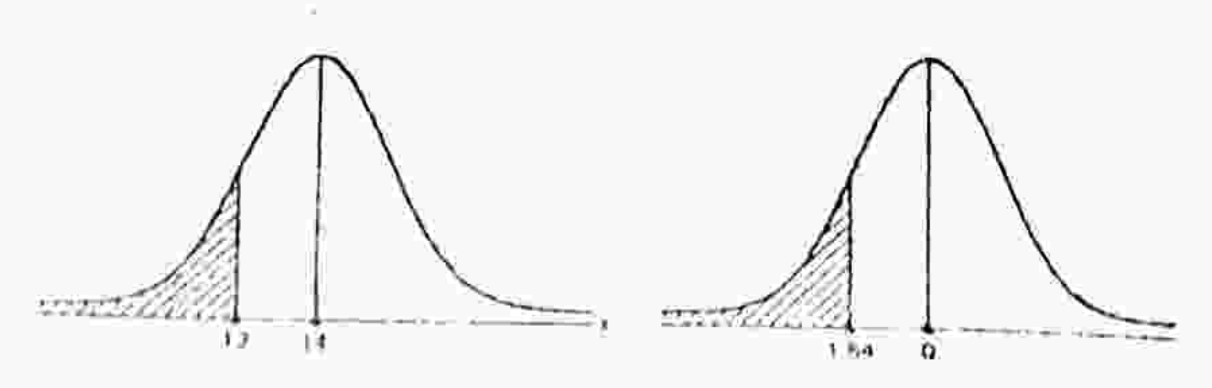
بفرض أن المتغير العشوائي X يمثل أوزان الخراف بهذا المشروع وبالتالي فإن :

$$X \sim N(\mu = 14 \ , \ \sigma^2 = 1.4884)$$

وعليه فإن نسبة الخراف التي يقل وزنها عن 12 كجم يمثلها الاحتمال التالي : $P(X < 12)$ والذي يمكن إيجاده باستخدام التوزيع الطبيعي المعياري على النحو الآتي :

$$P(X < 12) = P\left(Z < \frac{12 - \mu}{\sigma}\right) = P\left(Z < \frac{12 - 14}{1.22}\right) = P(Z < -1.64)$$

$$= 0.5 - P(-1.64 \leq Z < 0) = 0.5 - 0.4495 = 0.0505$$



شكل (28) : المساحة المظللة = $P(X < 12) = P(Z < -1.64)$

مثال (31): إذا كان المتغير العشوائي X يتوزع وفق التوزيع الطبيعي بمتوسط 25.5 وتباين مجهول وعلمت أن $P(X < 27.8) = 0.7549$ فما قيمة الانحراف المعياري لذلك المجتمع ؟

الحل :

$$P(X < 27.8) = 0.7549 \quad \Leftrightarrow \quad P\left(Z < \frac{27.8 - \mu}{\sigma}\right) = 0.7549$$

$$\Leftrightarrow \quad P\left(Z < \frac{27.8 - 25.5}{\sigma}\right) = 0.7549 \quad \Leftrightarrow \quad P\left(Z < \frac{2.3}{\sigma}\right) = 0.7549$$

$$\Leftrightarrow \quad P(Z < z_0) = 0.7549$$

حيث

$$z_0 = \frac{2.3}{\sigma}$$

والذي يمكن إعادة كتابته كما يلي :

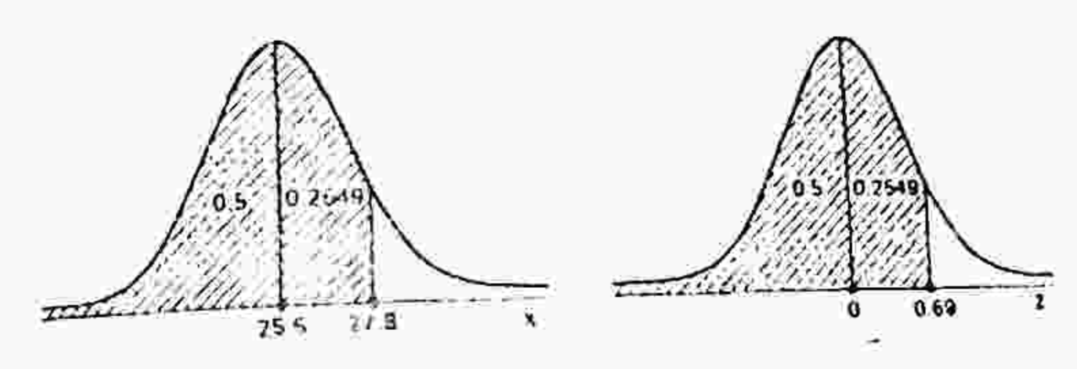
$$P(Z \le z_0) = P(-\infty < Z \le 0) + P(0 \le Z \le z_0)$$

$$\Rightarrow \quad 0.7549 = 0.5 + P(0 \le Z \le z_0)$$

$$\Rightarrow \quad P(0 \le Z \le z_0) = 0.2549 \quad \Rightarrow \quad z_0 = 0.69$$

وذلك من جدول التوزيع الطبيعي المعياري . وبالتالي فإن :

$$z_0 = \frac{2.3}{\sigma} = 0.69 \quad \Rightarrow \quad \sigma = \frac{2.3}{0.69} = 3.33$$



شكل (29) : المساحة المضللة = $P(X < 27.8) = P(Z < 0.69)$

مثال (32) : إذا كانت أوزان أرغفة الخبز التي ينتجها أحد المخابز تتبع في تغيراتها التوزيع الطبيعي بمتوسط 0.5 وحدة وزن وانحراف معياري 0.01 وحدة وزن ، فإذا علمت أن نسبة أوزان الأرغفة التي يقل وزنها عن وزن معين وليكن x_0 تساوي 0.8 ٪ فما هو ذلك الوزن ؟

الحـل :

بفرض أن المتغير العشوائي X يمثل أوزان أرغفة الخبز بهذا المخبز وبالتالي فإن :

$$X \sim N(\mu = 0.5 \ , \ \sigma^2 = 0.0001)$$

وبذلك يكون المطلوب إيجاد قيمة x_0 التي تحقق الشرط التالي : $P(X < x_0) = 0.008$

وعليه فإن :

$$P(X < x_0) = 0.008 \quad \Leftrightarrow \quad P\left(Z < \frac{x_0 - \mu}{\sigma}\right) = 0.008$$

$$\Leftrightarrow \quad P\left(Z < \frac{x_0 - 0.5}{0.01}\right) = 0.008 \quad \Leftrightarrow \quad P(Z < z_0) = 0.008$$

$$P(-\infty < Z < z_0) = 0.008$$

أي أن :

حيث

$$z_0 = \frac{x_0 - 0.5}{0.01}$$

(أ)

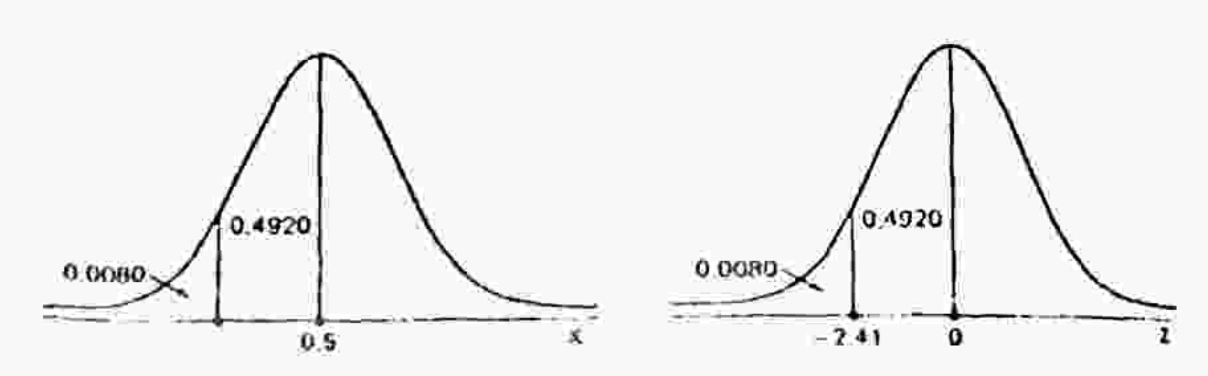
ومن جدول التوزيع الطبيعي المعياري نجد أن :

$$P(Z < z_0) = 0.008 \quad \Rightarrow \quad z_0 = -2.41$$

(ب)

من (أ) و (ب) نستنتج أن :

$$z_0 = \frac{x_0 - 0.5}{0.01} = -2.41 \quad \Rightarrow \quad x_0 = 0.5 + 0.01(-2.41) = 0.5 - 0.0241 = 0.4759$$



شكل (30) : المساحة المظللة = $P(X < 0.4759) = P(Z < -2.41)$.

مثال (33) : بفرض أن درجات الطلبة في مادة الإحصاء تتوزع وفق التوزيع الطبيعي بمتوسط 76 وانحراف معياري 4 فإذا كانت أقل درجة بدرجات تقدير مقبول (D) تساوي 58 وبدرجات تقدير جيد (C) تساوي 65 ، وبدرجات تقدير جيد جداً (B) تساوي 82 ، وبدرجات تقدير ممتاز (A) تساوي 88 ، فما هي نسبة الطلبة الذين سيحصلون على كل تقدير من التقديرات السابقة ؟

الحل :
بفرض أن المتغير العشوائي X يمثل درجات الطلبة بهذا المقرر فإن $X \sim N(76 , 16)$
وعليه فإن :
نسبة الطلبة الذين تقديراتهم ممتاز (A) هي :

$$P(X \geq 88) = P\left(\frac{X-76}{4} \geq \frac{88-76}{4}\right) = P(Z \geq 3) = 0.0013$$

نسبة الطلبة الذين تقديراتهم جيد جدا (B) هي:

$$\begin{aligned} P(82 \leq X \leq 88) &= P(1.5 \leq Z < 3) \\ &= P(Z<3) - P(Z<1.5) \\ &= 0.9987 - 0.9332 = 0.0655 \end{aligned}$$

نسبة الطلبة الذين تقديراتهم جيد (C) هي :

$$\begin{aligned} P(66 \leq X < 82) &= P(-2.5 \leq Z < 1.5) \\ &= P(Z<1.5) - P(Z<-2.5) \\ &= 0.9332 - 0.0062 = 0.9270 \end{aligned}$$

نسبة الطلبة الذين تقديراتهم مقبول (D) هي :

$$\begin{aligned} P(58 \leq X < 66) &= P(-4.5 \leq Z < 2.5) \\ &= P(Z \leq -2.5) - P(Z<-4.5) \\ &= 0.0062 - 0.0000 = 0.0062 \end{aligned}$$

مما سبق يتضح جلياً بانه لا يوجد راسبين في هذه المادة وذلك لأن مجموع هذه النسب يساوي واحد .

5 - 3 - 5 تقريب توزيع ذي الحدين باستخدام التوزيع الطبيعي
Normal Approximation to the Binomial

لقد تعرضنا في بند 5 - 2 - 3 لتوزيع ذي الحدين ، وأشرنا لكيفية حساب الاحتمالات باستخدام دالة كتلة احتماله ، ولكن في العديد من المسائل التي يكون فيها عدد المحاولات (n) كبيراً نجد أن طريقة حساب الاحتمالات باستخدام دالة الكتلة الاحتمالية معقدة وغير عملية، ولتفادي هذه الصعوبة يمكن استخدام التوزيع الطبيعي للحصول على قيم تقريبية لتلك الاحتمالات حيث يكون هذا التقريب جيداً عندما تكون $n > 20$ و $n\,p \geq 5$ وبالتالي فإن $n\,q > 5$. أما .

كانت قيمة p قريبة من الصفر أو الواحد الصحيح فإنه يكون غير جيد مهما كانت قيمة n .
والنظرية التالية تتضمن هذا التقريب .

نظرية (13) : إذا كان المتغير العشوائي X يتوزع وفق توزيع ذي الحدين بمعلمتين n و p ، وكانت $n > 20$ و $n\,p \geq 5$ وكان المتغير العشوائي Y يتوزع وفق التوزيع الطبيعي بمتوسط $\mu = np$ وتباين $\sigma^2 = npq$ ، فإنه لأى عددين صحيحين a و b يكون :

$$P(a \leq X \leq b) \approx P(a - 0.5 \leq Y \leq b + 0.5)$$

$$= P\left(\frac{(a-0.5)-n\,p}{\sqrt{n\,p\,q}} \leq \frac{Y - n\,p}{\sqrt{n\,p\,q}} \leq \frac{(b+0.5)-n\,p}{\sqrt{n\,p\,q}}\right)$$

$$= P\left(\frac{(a-0.5)-n\,p}{\sqrt{n\,p\,q}} \leq Z \leq \frac{(b+0.5)-n\,p}{\sqrt{n\,p\,q}}\right)$$

$$= \Phi\left(\frac{(b+0.5)-n\,p}{\sqrt{n\,p\,q}}\right) - \Phi\left(\frac{(a-0.5)-n\,p}{\sqrt{n\,p\,q}}\right) \qquad (57)$$

حيث $\Phi(.)$ تمثل دالة التوزيع التراكمي للتوزيع الطبيعي المعياري .

الحظ أن التوزيع الطبيعي توزيع متصل بينما توزيع ذي الحدين توزيع منفصل ، لذلك وجد العالم ياتس (Yates) أن التقريب يكون اكثر دقة وواقعية إذا أجرينا تصحيحاً على قيمة x وذلك إما بإضافة أو طرح 0.5 منها ، ويسمى هذا التصحيح أحياناً بتصحيح ياتس نسبة للعالم ياتس أو معامل التصحيح أو تصحيح الاتصال (Continuity Correction) .
وعليه فإنه من المألوف من الناحية العملية استخدام معامل التصحيح (0.5) الذي في الواقع يعد ضرورياً عند حساب ($P(X = x)$. والمثال التالي يوضح كيفية استخدام هذه النظرية .

مثال (34): بفرض أنه ألقيت قطعة نقود معدنية خمسة عشرة مرة ، ولنفرض أن احتمال ظهور صورة (H) يساوي 0.4 وأن المتغير العشوائي X يمثل عدد الصور التي يمكن

الحصول عليها ، احسب الاحتمالين التاليين $P(X=4)$ و $P(7 \le X \le 9)$ باستخدام دالة كتلة احتمال توزيع ذي الحدين ثم باستخدام التقريب بالتوزيع الطبيعي .

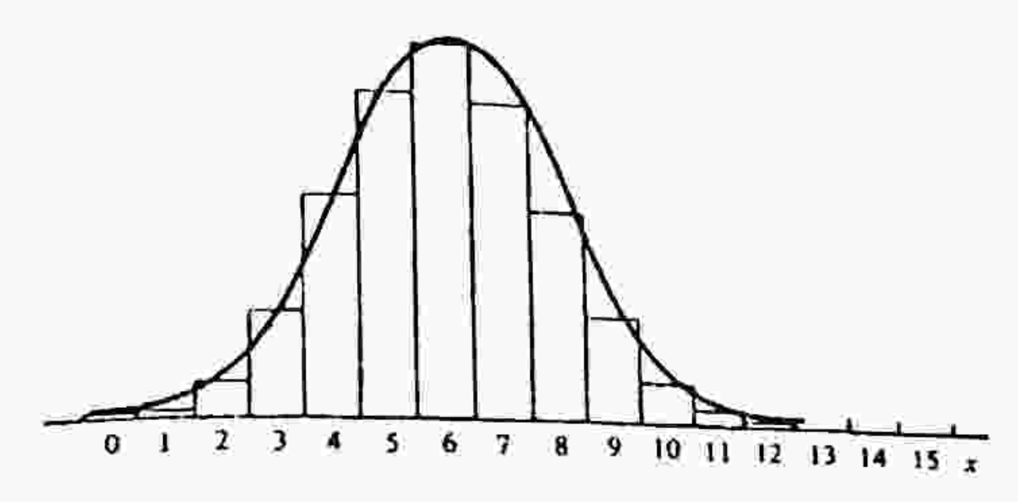
الحل :

حيث أن المتغير العشوائي X يتوزع وفق توزيع ذي الحدين بمعلمتين $p = 0.4$ و $n = 15$

وعليه فإن دالة كتلة احتماله تكون على الصيغة التالية :

$$p_X(x) = P(X=x) = \begin{cases} \binom{15}{x}(0.4)^x(0.6)^{15-x} & , x = 0,1,2,\dots,15 \\ 0 & \text{خلاف ذلك ,} \end{cases}$$

والشكل (31) يوضح التمثيل البياني لهذه الدالة .



شكل (31) : دالة كتلة الاحتمال لتوزيع ذي الحدين (B (n = 15 , p = 0.4

وتقريبها بالتوزيع الطبيعي .

وعليه فإن :

$$P(X=4) = \binom{15}{4}(0.4)^4(0.6)^{11} = 01268$$

وإن :

$$P(7 \le X \le 9) = \sum_{x=7}^{9} \binom{15}{x} (0.4)^x (0.6)^{15-x}$$

$$= P(X \le 9) - P(X \le 6)$$

$$= \sum_{x=0}^{9} \binom{15}{x} (0.4)^x (0.6)^{15-x} - \sum_{x=0}^{6} \binom{15}{x} (0.4)^x (0.6)^{15-x}$$

$$= 0.9662 - 0.6098 = 0.3564$$

وباستخدام التقريب الطبيعي الحظ أن :

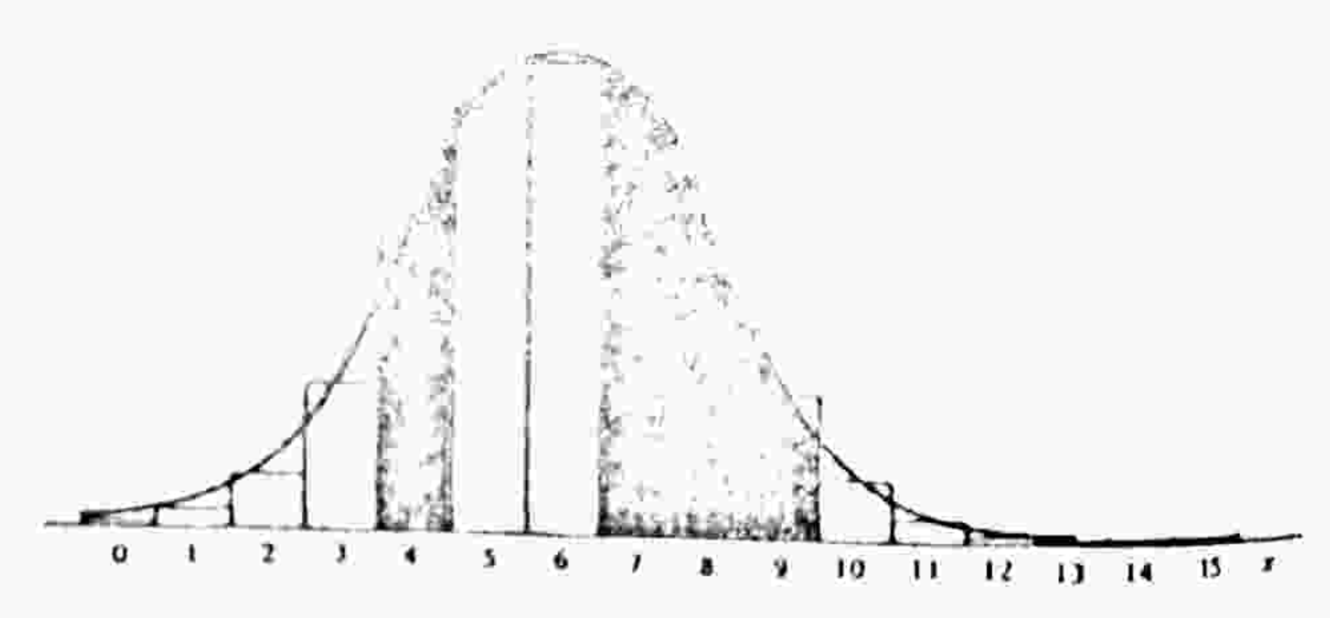
$$\mu = n\,p = (15)(0.4) = 6$$

وأن

$$\sigma^2 = n\,p\,q = (15)(0.4)(0.6) = 3.6$$

وبرسم منحنى التوزيع الطبيعي $N(6\ ,\ 3.6)$ على نفس الشكل (31) الذي يمثل دالة كتلة الاحتمال لتوزيع ذي الحدين $B(n = 15\ ,\ p = 0.4)$ ، نلاحظ أن $P(X = 4)$ باستخدام توزيع ذي الحدين يساوي مساحة المستطيل الذي قاعدته مركزها عند $x = 4$.

بينما باستخدام التقريب فإن هذا الاحتمال يساوي المساحة المضللة تحت المنحنى الطبيعي الواقعة بين الإحداثيات $x_1 = 3.5$ و $x_2 = 4.5$ كما في شكل (32) أدناه :



شكل (32) : التقريب الطبيعي لتوزيع ذي الحدين

وبالتحويل إلى التوزيع الطبيعي المعياري نجد أن :

$$z_1 = \frac{3.5-6}{1.9} = -1.316 \quad و \quad z_2 = \frac{4.5-6}{1.9} = -0.789$$

وعليه فإن :

$$P(X=4) \cong P(-1.316 < Z < -0.789)$$
$$= P(Z < -0.789) - P(Z < -1.316)$$
$$= 0.2115 - 0.0941 = 0.1210$$

وهذا يتفق إلى حد كبير مع قيمة الاحتمال الفعلية 0.1268 التي تم حسابها باستخدام توزيع ذي الحدين . بينما الاحتمال $P(7 \le X \le 9)$ يساوي مجموع مساحة المستطيلات التي مراكز قواعدها عند x = 7 ، 8 ، 9 . وباستخدام التقريب الطبيعي نجد أن تلك المساحة تساوي تقريباً مساحة المنطقة المضللة تحت المنحنى الطبيعي بين $x_1 = 6.5$ و $x_2 = 9.5$ وذلك كما هو موضح في شكل (32) . إن قيم Z المناظرة هي :

$$z_1 = \frac{6.5-6}{1.9} = 0.263 \quad , \quad z_2 = \frac{9.5-6}{1.9} = 1.482$$

وعليه فإن :

$$P(7 \le X \le 9) \cong P(0.263 \le Z \le 1.842)$$
$$= P(Z \le 1.842) - P(Z \le 0.263)$$
$$= \Phi(1.842) - \Phi(0.263) = 0.9673 - 0.6037 = 0.3636$$

مرة أخرى نلحظ أن قيمة الاحتمال باستخدام التقريب الطبيعي تتفق إلى حد ما من القيمة الحقيقية 0.3564 التي تم الحصول عليها من توزيع ذي الحدين .

مثال (35): إذا كان $X \sim B(n=100 , p=0.2)$ احسب الاحتمالين التاليين باستخدام التوزيع الطبيعي :
أ - $P(X < 18)$ ب - $P(X \ge 22)$

الحل :

حيث $\mu = np = 20$ و $\sigma = \sqrt{npq} = 4$ فإن :

أ-

$$P(X<18) = P(X \le 17) \cong P(X + 0.5 \le 17.5) = P\left(Z \le \frac{17.5-20}{4}\right)$$
$$= P(Z \le -0.625) = \Phi(-0.625) = 0.266$$

ب - وبالمثل

$$P(X\geq 22)=1-P(X<22)=1-P(X\leq 21)\equiv 1-P(X+0.5\leq 21.5)$$

$$=1-P\left(Z\leq\frac{21.5-20}{4}\right)=1-P(Z\leq 0$$

$$=1-\Phi(0.375)=1-0.6459=0.354$$

من خـلال النظريـة (13) والمثـالين (34) و (35) يمكـن تلخيـص الصيـغ المختلفـة احتمالات توزيع ذي الحدين باستخدام التقريب الطبيعي في الجدول التالي :

جدول (2) : صيغ حساب احتمالات توزيع ذي الحدين باستخدام التوزيع الطبيعي .

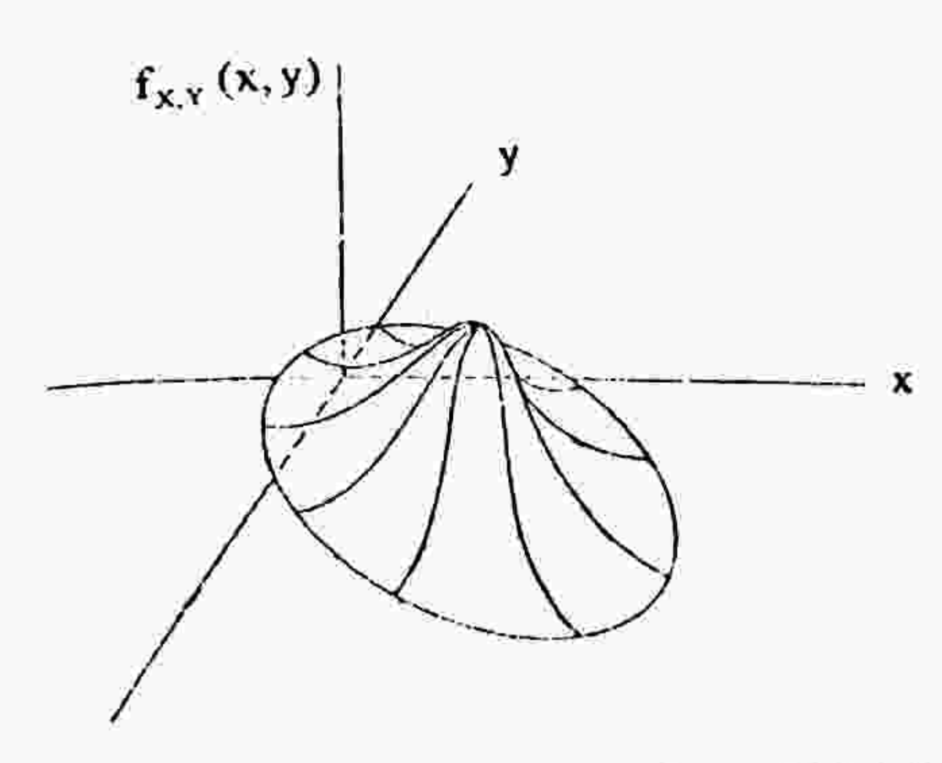
احتمالات ذي الحدين	التقريب الطبيعي
$P(X=a)$	$\Phi\left(\frac{(a+0.5)-np}{\sqrt{npq}}\right)-\Phi\left(\frac{(a-0.5)-np}{\sqrt{npq}}\right)$
$P(X\leq a)$	$\Phi\left(\frac{(a+0.5)-np}{\sqrt{npq}}\right)$
$P(X<a)$	$\Phi\left(\frac{(a-0.5)-np}{\sqrt{npq}}\right)$
$P(X\geq -a)$	$1-\Phi\left(\frac{(a-0.5)-np}{\sqrt{npq}}\right)$
$P(X>a)$	$1-\Phi\left(\frac{(a+0.5)-np}{\sqrt{npq}}\right)$

6 - 3 التوزيع الطبيعي الثنائي Bivariate Normal Distribution

يُعد هذا التوزيع من أهم التوزيعات الثنائية التي تستخدم لتحليل التغاير Covariance الارتباط Correlation . فإذا كان X و Y متغيرين عشوائيين بدالة كثافة احتمال مشتركة لها صيغة الآتية :

$$f_{X,Y}(x,y)=\frac{1}{2\rho\sigma_1\sigma_2\sqrt{1-\rho^2}}e^{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_1}{\sigma_1}\right)^2-2\rho\left(\frac{x-\mu_1}{\sigma_1}\right)\left(\frac{y-\mu_2}{\sigma_2}\right)+\left(\frac{y-\mu_2}{\sigma_2}\right)^2\right]} \quad (58)$$

حيث $-\infty < x < \infty$ ، $-\infty < y < \infty$ ، $-\infty < \mu_1 < \infty$ ، $-\infty < \mu_2 < \infty$ ، $\sigma_1 > 0$
$\sigma_2 > 0$ ، $|\rho| < 1$. فإن المتغيرين X و Y يتوزعان وفق التوزيع الطبيعي الثنائي
الذي يرمز له بالرمز : $(X\ ,\ Y) \sim BVN\left(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho\right)$. والشكل (33)
أدناه يوضح الرسم البياني لدالة كثافة الاحتمال المشتركة للمتغيرين X و Y .



شكل (33) : دالة كثافة الاحتمال للتوزيع الطبيعي الثنائي

وإذا كانت $\mu_1 = \mu_2$ و $\sigma_1 = \sigma_2 = 1$ فإن $(X, Y) \sim BVN(0, 0, 1, 1, \rho)$
أي أن :

$$f_{X,Y}(x,y) = \frac{1}{2\pi} e^{-\frac{1}{2(1-\rho^2)}\left(x^2 - 2\rho x y + y^2\right)} \quad -\infty < x < \infty\ ,\ -\infty < y < \infty\ ,\ |\rho| < 1 \qquad (59)$$

وإن دالة كثافة الاحتمال الهامشية للمتغير X تكون كالآتي :

$$\begin{aligned} f_X(x) &= \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy \\ &= \frac{1}{\sqrt{2\pi\sigma_1^2}} e^{-\frac{1}{2\sigma_1^2}(x-\mu_1)^2} \quad , -\infty < x < \infty \end{aligned} \qquad (60)$$

$$X \sim N\left(\mu_1\ ,\ \sigma_1^2\right)$$

أي أن
وبالمثل دالة كثافة الاحتمال الهامشية للمتغير العشوائي Y تكون كالآتي :

$$f_Y(y)=\frac{1}{\sqrt{2\pi\sigma_2^2}}\, e^{-\frac{1}{2\sigma_2^2}(y-\mu_2)^2} \qquad ,\ -\infty < y < \infty \tag{61}$$

$$Y \sim N\left(\mu_2\ ,\ \sigma_2^2\right)$$

أي أن
دالة كثافة الاحتمال الشرطية للمتغير العشوائي X بشرط $Y = y$ تكون كالتالي :

$$g_1(x|y)=\frac{f_{X,Y}(x,y)}{f_Y(y)}$$

$$=\frac{1}{\sigma_1\sqrt{2\pi}\sqrt{1-\rho^2}}\, e^{-\frac{1}{2(1-\rho^2)\sigma_1^2}\left[x-\left(\mu_1+\rho\frac{\sigma_1}{\sigma_2}(y-\mu_2)\right)\right]^2} \qquad ,\ -\infty < x < \infty$$

وعليه فإن

$$g_1(X|Y=y) \sim N\left(\mu_1+\rho\frac{\sigma_1}{\sigma_2}(y-\mu_2)\ ,\ \sigma_1^2\left(1-\rho^2\right)\right) \tag{62}$$

أي أن :

$$E(X|y)=\mu_1+\rho\frac{\sigma_1}{\sigma_2}(y-\mu_2) \tag{63}$$

$$V(X|y)=\sigma_1^2\left(1-\rho^2\right) \tag{64}$$

كذلك فإن دالة كثافة الاحتمال الشرطية للمتغير العشوائي Y بشرط $X = x$ تأخذ الصيغة الآتية :

$$g_2(Y|X=x) \sim N\left(\mu_2+\rho\frac{\sigma_2}{\sigma_1}(x-\mu_1)\ ,\ \sigma_2^2\left(1-\rho^2\right)\right) \tag{65}$$

أي أن :

$$V(Y|x)=\sigma_2^2\left(1-\rho^2\right) \qquad , \qquad E(Y|x)=\mu_2+\rho\frac{\sigma_2}{\sigma_1}(x-\mu_1)$$

وأن $\operatorname{cov}(X,Y)=\rho\sigma_1\sigma_2$ لذلك فإن معامل الارتباط بين X و Y هو ρ ، وفي الحقيقة أن :

$$E(XY)=E\left[E(XY|X)\right]$$

حيث

$$E(XY|x)=x\,E(Y|x)=x\mu_2+\rho\frac{\sigma_2}{\sigma_1}\,.x(x-\mu_1)$$

وعليه فإن :

$$E[E(XY|X)] = E\left[X\mu_2 + \rho\frac{\sigma_2}{\sigma_1}X(X-\mu_1)\right]$$

$$= \mu_1\mu_2 + \rho\frac{\sigma_2}{\sigma_1}\left[E(X^2) - \mu_1^2\right]$$

$$= \mu_1\mu_2 + \rho\frac{\sigma_2}{\sigma_1}\cdot\sigma_1^2 = \mu_1\mu_2 + \rho\sigma_1\sigma_2$$

إذن

$$E(XY) = \mu_1\mu_2 + \rho\sigma_1\sigma_2$$

وبالتالي فإن :

$$\operatorname{cov}(X,Y) = E(XY) - E(X)E(Y) = \rho\sigma_1\sigma_2 \qquad (66)$$

ملحوظة :

1 - إذا كان المتغيران العشوائيان X و Y يتوزعان وفق التوزيع الطبيعي الثنائي فإن X و Y مستقلان إذا وفقط إذا كان $\rho = 0$.

2 - إذا كان X و Y متغيرين عشوائيين يتبعان التوزيع الطبيعي الثنائي وكانت a_1, a_2, b أعداداً حقيقية فإن المتغير W حيث $W = a_1X + a_2Y + b$ ، يتبع التوزيع الطبيعي بمتوسط يساوي $E(W) = a_1\mu_1 + a_2\mu_2 + b$ وتباين يساوي

$$V(W) = a_1^2\sigma_1^2 + a_2^2\sigma_2^2 + 2a_1a_2\rho\sigma_1\sigma_2$$

مثال (36) : إذا تم اختيار امرأة ورجل من مجتمع المتزوجين بطريقة عشوائية وكان التوزيع المشترك لطول الرجل والمرأة بذلك المجتمع يتبع التوزيع الطبيعي الثنائي ، وكان متوسط أطوال النساء يساوي 156.8 سم وانحرافها المعياري 2 سم ، بينما متوسط أطوال الرجال يساوي 160 سم وانحرافها المعياري 2 سم ، وأن معامل الارتباط بين الطولين يساوي 0.68 فما احتمال أن تكون أطوال النساء اكبر من أطوال الرجال بهذا المجتمع ؟

الحل : لنفرض أن المتغير العشوائي X يمثل أطوال النساء والمتغير العشوائي Y يمثل أطوال الرجال ، إذن الاحتمال المطلوب هو $P(\,(X - Y) > 0\,)$. وحيث أن X و Y لهما توزيع طبيعي ثنائي وبالتالي فان $X - Y$ يكون له توزيع طبيعي بمتوسط :

$$E(X - Y) = E(X) - E(Y) = 156.8 - 160 = -3.2$$

وبتباين

$$\begin{aligned} V(X - Y) &= V(X) + V(Y) - 2\,Cov(X, Y) \\ &= 4 + 4 - 2(0.68)(2)(2) = 2.56 \end{aligned}$$

وبانحراف المعياري يساوي 1.6 إذن المتغير العشوائي Z يتوزع وفق التوزيع الطبيعي المعياري حيث $Z = \dfrac{(X - Y) + 3.2}{1.6}$ أي أن $Z \sim N(0\ ,\ 1)$. ومن جدول التوزيع الطبيعي نجد أن

$$P(\,(X - Y) > 0\,) = P(Z > 2) = 1 - \Phi(2) = 0.0227$$

5 - 3 - 7 توزيع مربع كاي The Chi - Square Distribution

إن لهذا التوزيع علاقة بالتوزيع الطبيعي وله تطبيقات عديدة في مجال الإحصاء الاستنتاجي نذكر منها على سبيل المثال : اختبارات جودة المطابقة والتجانس والاستقلالية والتباين .... الخ . يقال أن المتغير العشوائي X يتوزع وفق توزيع مربع كاي بدرجات حرية تساوي n إذا كانت دالة كثافة احتماله لها الصيغة الآتية :

$$f_X(x) = \begin{cases} \dfrac{1}{2^{n/2}\,\Gamma(n/2)}\, x^{\frac{n}{2}-1}\, e^{-x/2} & , x > 0 \\ 0 & , x \le 0 \end{cases} \qquad (67)$$

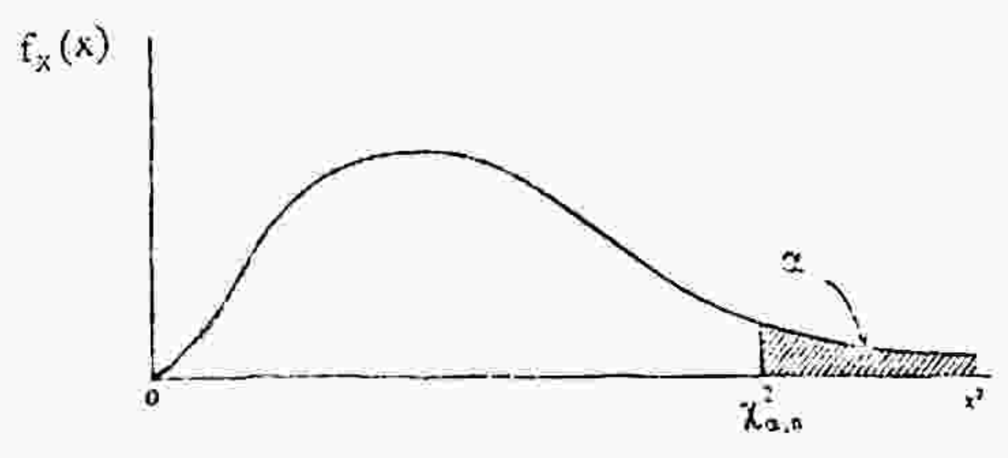
ويرمز لذلك عادة بالرمز $X \sim \chi^2_{(n)}$. وإن :

$$E(X)=n$$
$$V(X)=2n$$
$$m_X(t)=\left(\frac{1}{1-2t}\right)^{n/2} \quad , t<1/2 \qquad (68)$$

وحيث إن دالة كثافة هذا التوزيع تعتمد على n وبالتالي لكل قيمة من قيم n سيكون لهذه الدالة شكل خاص بها . ولهذا السبب وضعت جداول خاصة بهذا التوزيع لقيم n = 1 ، 2 ، 3 ، . . . وذلك لحساب الاحتمالات تحت منحنى هذا التوزيع، وهي معطاة بجدول (6) في آخر الكتاب وبتعريف $\chi^2_{\alpha,n}$ القيمة المئوية لمتغير مربع كاى بدرجات حرية تساوي n بحيث أن احتمال أن $\chi^2_{(n)}$ أكبر من هذه القيمة يساوي α ، أي أن :

$$P\left(\chi^2_{(n)} > \chi^2_{\alpha,n}\right)=\int_{\chi^2_{\alpha,n}}^{\infty} f_X(x)\,dx=\alpha$$

هذا الاحتمال موضح في المنطقة المظللة في الشكل التالي :



شـكل (34) : القيمة المئوية $\chi^2_{\alpha,n}$ لتوزيع مربع كاى .

ومن شكل (34) يتضح أن هذا التوزيع ملتوى ناحية اليمين وله منوال واحد .

مثال (37) : إذا كانت $\chi^2_{(n)}$ ترمز لمتغير عشوائي يتبع توزيع مربع كاى بدرجات حرية n من جدول (6) أوجد : أ - $\chi^2_{0.01,5}$ ب - $\chi^2_{0.05,20}$ جـ - $\chi^2_{0.90,28}$.

حل :

من جدول (6) و بالنظر في العمود n عندما n = 5 و الصف α عندما α = 0.01 نجد التقاطع بينهما عند القيمة 15.076 أي أن $\chi^2_{0.01,5} = 15.076$.

من جدول (6) و بالنظر في العمود n و عندما n = 20 و الصف α و عندما α = 0.05 نجد أن التقاطع بينهما عند القيمة 31.410 أي أن $\chi^2_{0.05,20} = 31.410$.

من جدول (6) و بالنظر في عمود n و عندما n = 28 و الصف α عندما α = 0.90 نجد أن التقاطع بينهما عند القيمة 18.939 نجد أن $\chi^2_{0.90,28} = 18.939$.

مثال (38) : إذا كانت $\chi^2_{(n)}$ ترمز لمتغير عشوائي يتبع توزيع كاى بدرجات حرية تساوى n

من جدول مربع كاى أوجد :

أ - $P\left(\chi^2_{(13)} < 9.926\right)$ ب - $P\left(\chi^2_{(26)} \le -10\right)$

ج - $P\left(\chi^2_{(25)} \ge 34.382\right)$ د - $P\left(9.591 \le \chi^2_{(20)} \le 34.178\right)$

الحل :

حيث أن جدول (6) يعطى القيم المئوية العليا و عليه فإن :

$$P\left(\chi^2_{(13)} < 9.926\right) = 1 - P\left(\chi^2_{(13)} \ge 9.926\right)$$

بالنظر في العمود n عندما n = 13 نجد أن العدد 9.926 يقع في العمود 0.70 و عليه فإن:

$$P\left(\chi^2_{(13)} < 9.926\right) = 1 - 0.70 = 0.30$$

- $P\left(\chi^2_{(26)} \le -10\right) = 0$ وذلك لأن متغير مربع كاى يأخذ قيم غير سالبة .

- من جدول (6) و بالنظر في العمود n عندما n = 25 نجد أن العدد 34.382 يقع في 0.10 و عليه فإن : $P\left(\chi^2_{(25)} \ge 34.382\right) = 0.10$.

د - من جدول (6) نجد أن :

$$P\left(9.591 \le \chi^2_{(20)} \le 34.170\right) = P\left(\chi^2_{(20)} \ge 9.591\right) - P\left(\chi^2_{(20)} \ge 34.170\right)$$
$$= 0.975 - 0.025 = 0.95$$

نظريــة (14) : إذا كــانت $X_1, X_2, \ldots, X_n$ متغـيرات عشـوائية مسـتقلة وكـان $X_i \sim N\left(\mu_i, \sigma_i^2\right)$ حيث $i = 1, 2, \ldots, n$ فإن المتغير العشوائي

$$Y = \sum_{i=1}^{n} \left(\frac{X_i - \mu_i}{\sigma_i} \right)^2$$

يتوزع وفق توزيع مربع كاي بدرجات حرية تساوي n .

نظريـة (15) : إذا كانت $X_1, X_2, \ldots, X_k$ متغيرات عشوائية مستقلة وكان $X_i \sim \chi^2_{(n_i)}$ حيث $i = 1, 2, \ldots, k$ فإن المتغير العشوائي $Y = \sum_{i=1}^{n} X_i$ يتوزع وفق توزيع مربع كاي بدرجات حرية تساوي $\sum_{i=1}^{k} n_i$.

ملحوظــة (3) : إذا كان المتغير العشوائي X يتبع توزيع مربع كاي بدرجـات حريـة n ، أي أن $X \sim \chi^2_{(n)}$ فإن المتغير العشوائي :

$$Z = \frac{X - n}{\sqrt{2n}}$$

يتوزع وفق التوزيع الطبيعي المعياري ، أي أن $X \sim N(0, 1)$ عندما $n \to \infty$.

5 - 3 - 8 توزيع t The t Distribution

في هذا البند سنناقش توزيعاً آخر يطلق عليه توزيع t وهو ذو علاقة بالعينات العشوائية التي يتم اختيارها من مجتمع طبيعي ، ولهذا التوزيع تطبيقات عديدة في مجال الإحصاء الاستنتاجي ، ولقد وضعت جداول خاصة به ويعرف هذا التوزيع كما يلي :

إذا كان X متغيراً عشوائياً يتبع التوزيع الطبيعي المعياري N(0,1) وكان Y متغيراً عشوائياً له توزيع مربع كاي بدرجات حرية n وكان المتغير العشوائي Y مستقلاً عن المتغير العشوائي X

إن المتغير العشوائي $T = \frac{X}{\sqrt{Y/n}}$ له توزيع يطلق عليه تسمية توزيع t بدرجات حرية تساوي n ، ويرمز لذلك بالرمز بالرمز $T \sim t_{(n)}$ ، وتكون دالة كثافة احتمال على الصورة التالية :

$$f_T(t) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{\pi n}\ \Gamma(1/2)} \left(1 + \frac{t^2}{n}\right)^{-(n+1)/2} \quad , \ t \in R \qquad (69)$$

بعض سمات توزيع t :

لتوزيع t العديد من الخواص الهامة من بينها :

1- إن $f_T(t)$ متماثلة أي أن $f_T(t) = f_T(-t)$ ولهما شكل ناقوسي يشبه شكل التوزيع الطبيعي وأن $f_T(t) \to 0$ كلما اقتربت t من ∞ ، وعندما تكون n كبيرة فإن توزيع t يقترب من التوزيع الطبيعي ، أي أن عندما $n \to \infty$ يكون :

$$f_T(t) \to \frac{1}{\sqrt{2\pi}} e^{-t^2/2}$$

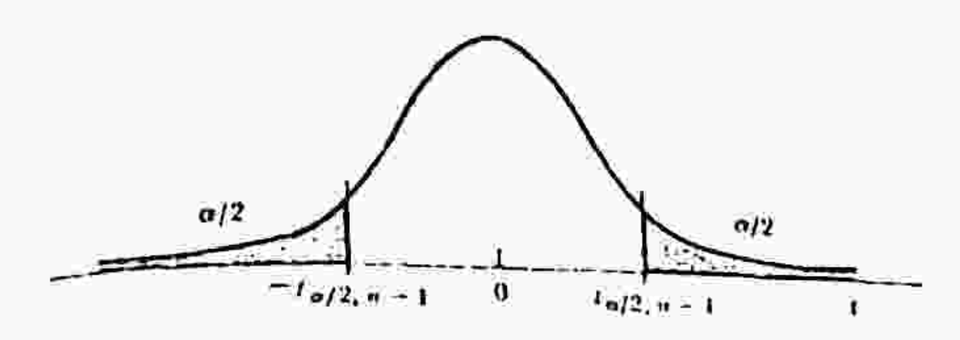
ولكن عندما تكون n صغيرة فإن توزيع t يختلف بشكل واضح عن التوزيع الطبيعي . وفي الواقع نجد أن :

$$P(|T| \ge t_0) \ge P(|Z| > t_0) \qquad , Z \sim N(0, 1)$$

وهذا يعني أن هناك احتمال اكبر في طرف توزيع t مقارنة بالتوزيع الطبيعي المعياري . وسوف نستخدم الرمز $t_{\alpha,n}$ لقيم T حيث :

$$P(T \ge t_{\alpha,n}) = \int_{t_{\alpha,n}}^{\infty} f_T(t)\,dt = \alpha$$

وجدول (5) في آخر الكتاب يعطي قيم $t_{\alpha,n}$ وذلك عندما n = 1 ، 2 ، 3 ، .... ، 30 ، 60 حيث $t_{\alpha,n}$ تمثل التجزؤ ذا المرتبة $1-\alpha$. وحيث إن التوزيع متماثل وعليه فإن القيم السالبة يمكن الحصول عليها من خلال وضع $t_{1-\alpha,n} = -t_{\alpha,n}$. كما هو موضح في الشكل التالي :



شكل (35) : القيم المئوية لتوزيع t

2 - القيمة المتوقعة للمتغير العشوائي T تساوي صفر أي أن:

$$E(T)=0$$

و إذا كانت 2 < n فإن :

$$E(T^2)=V(T)=n(n-2)$$

مثال (39) : باستخدام جدول (5) أوجد :

ا - $t_{0.05,12}$ ب - $t_{0.025,20}$ جـ - $t_{0.925,15}$ د - $t_{0.90,18}$

الحل :

ا - حيث أن $t_{0.05,12}$ تترك مساحة قدرها 0.05 على الطرف الأيمن من التوزيع و بالتالي من جدول (5) نجد أن $t_{0.05,12}=1.782$.

ب - حيث أن $t_{0.025,20}$ تترك مساحة قدرها 0.025 على الطرف الأيمن من التوزيع و عليه من جدول (5) نجد أن $t_{0.025,20}=2.086$.

جـ - حيث أن $t_{1-\alpha,n}=-t_{\alpha,n}$ و عليه فإن : $t_{0.975,15}=-t_{0.025,15}=-2.131$.

د - بالمثل كما في الحالة (جـ) نجد أن $t_{0.90,18}=-t_{0.10,18}=-1.330$.

مثال (40) : أوجد :

أ - $P(-t_{0.005} < T < t_{0.025})$

ب - $P(-2.262 < T < a) = 0.925$ حيث a مقدار ثابت و بدرجات حرية تساوى 9 .

الحل :

أ - إن المساحة التي على يمين $t_{0.025}$ تساوى 0.025 و المساحة التي على يسار $-t_{.005}$ تساوى 0.005 و عليه فإن المساحة ما بين $-t_{.005}$ و $t_{0.025}$ تكون كالتالي :

$$P(-t_{.005} < T < t_{0.025}) = 1 - (0.025 + 0.005) = 0.97 \; .$$

ب - إن قيمة الثابت a يجب أن تكون موجبة ذلك لأن المساحة ما بين 2.262 - و a قريبة جداً من الواحد . ومن جدول (5) وبدرجات حرية تساوى 9 نجد أن $t_{0.025} = 2.262$ و عليه فإن $-t_{0.025} = -2.262$ وإن المساحة على يسار 2.262 - تساوى 0.025 ، وبالتالي فإن إجمالي المساحة على يسار a تساوى $0.025 + 0.925 = 0.95$ وإن المساحة على يمين a تساوى $1 - 0.95 = 0.05$ ، وعليه من جدول (5) نجد أن $a = t_{0.05,9} = 1.833$

5 - 3 - 9 توزيع F The F- Distribution

يُعد توزيع F من التوزيعات الإحصائية الهامة حيث يستخدم في الإحصاء الاستنتاجي لإجراء العديد من اختبارات الفروض المتعلقة بتحليل التباين وتصميم التجارب واختبار معنوية خطوط الانحدار ، إلى غير ذلك من التطبيقات الإحصائية العديدة والهامة . ولقد اكتشف هذا التوزيع من قبل العالم الإنجليزي الشهير فيشر (R . A . Fisher) حيث استخدمه لاختبار النسبة بين تبايني مجتمعين طبيعيين . ويعرض هذا التوزيع كما يلي :

إذا كان X و Y متغيرين عشوائيين ومستقلين وكان X له توزيع مربع كاى بدرجات حرية m والمتغير العشوائي Y له توزيع مربع كاى بدرجات حرية n ، فإن المتغير العشوائي :

$$F = \frac{X/m}{Y/n}$$

له توزيع f بدرجات حرية m و n ويرمز له بالرمز $F \sim f(m, n)$.ودالة كثافة احتماله لها الصيغة الآتية :

$$h_F(f)=\begin{cases}\dfrac{\Gamma\left(\dfrac{m+n}{2}\right)\left(\dfrac{m}{n}\right)^{\frac{m}{2}}}{\Gamma\left(\dfrac{m}{2}\right)\Gamma\left(\dfrac{n}{2}\right)}\cdot\dfrac{f^{\frac{m}{2}-1}}{\left[1+\dfrac{m}{n}f\right]^{\left(\frac{m+n}{2}\right)}} & , f<0 \\ 0 & , f\leq 0\end{cases} \qquad (70)$$

إن هذا التوزيع محدد بالكامل بالمعلمتين m و n .

ويعد هذا التوزيع ملتوياً التواء موجب (ناحية اليمين) ، وتقل درجة الالتواء كلما زادت درجات حرية البسط m أو المقام n أو كليهما معاً .

بعض سـمات توزيع f :

1 - إن $h_F(f)\to 0$ كلما $f\to\infty$ ، وإذا كانت $m>2$ فإن $0 \to (f)h_F$ كلما $f\to 0$.

2 - إذا كانت $m>2$ فإن دالة الكثافة وحيدة المنوال عند $f=\dfrac{n(m-2)}{m(n+2)}$ ، أما إذا كانت $m=2$ فإن المنوال يكون عند $f=0$ ، وإذا كانت $m=1$ فإن :

$h_F(f)\to\infty$ كلما $f\to 0$.

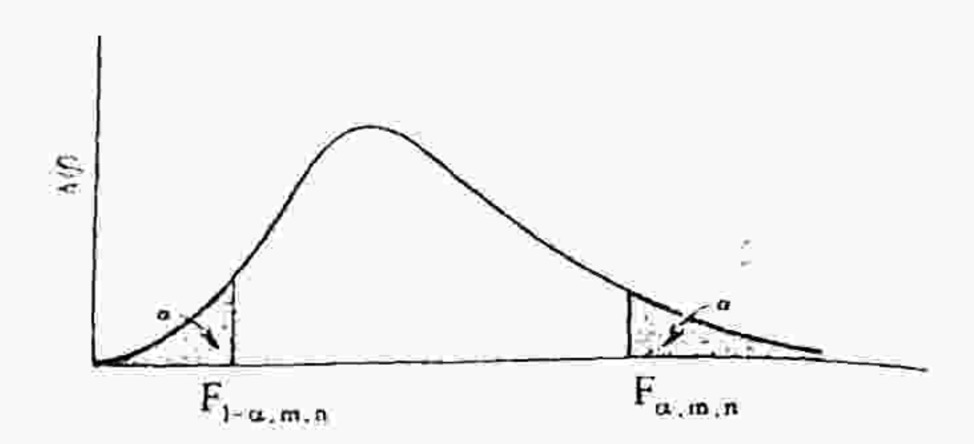
3 - إذا كان $X\sim f(m,n)$ فإن $\dfrac{1}{X}\sim f(n,m)$.

4 - إذا كانت $m=1$ فإن $f(1,n)=t^2_{(n)}$ ، أي أن $f(1,n)$ و T^2 لهما نفس التوزيع .

5 - سوف نرمز لقيم الطرف الأعلى من توزيع f بالرمز $F_{\alpha,m,n}$ وهذا يعني أن

$$P\left(F_{(m,n)}\geq f_{\alpha(m,n)}\right)=\int_{f_{\alpha,m,n}}^{\infty}h(f)\,df=\alpha$$

ونظراً لأهمية هذا التوزيع فقد وضعت جداول خاصة به عند قيم مختلفة لكل من α ، m ، n كما هو موضح في جدول (7) بآخر هذا الكتاب . والشكل التالي يوضح القيم المئوية العليا والسفلى لتوزيع F .



شكــل (36) : القيم المئوية العليا والسفلي لتوزيع F

وعلى ضوء الخاصية رقم (3) وحيث أن :

$$F_{m,n}=\frac{\chi^2_{(m)}/m}{\chi^2_{(n)}/n}=\frac{1}{\left(\chi^2_{(n)}/n\right)/\left(\chi^2_{(m)}/m\right)}=\left[F_{n,m}\right]^{-1}$$

وعليه فإن :

$$\alpha=P\left(F_{m,n}>f_{\alpha,m,n}\right)=P\left(f^{-1}_{\alpha,m,n}>F_{n,m}\right)=1-P\left(F_{n,m}>f^{-1}_{\alpha,m,n}\right)$$

وعليه فإن $f_{1-\alpha,n,m}=f^{-1}_{\alpha,m,n}$ ، وبالتالي من هذه الخاصية يمكن إيجاد القيم السفلى لمئويات توزيع F .

7. إذا كان $F\sim f_{(m,n)}$ فإن :

$$E(F)=\frac{n}{m}\cdot\frac{m/2}{(n/2)-1}=\frac{n}{n-2}\qquad ,\ n>2$$

و

$$V(F)=\frac{2n^2(m+n-2)}{m(n-2)^2(n-4)}\qquad ,\ n>4$$

مثال (**41**) : من جدول (7) أوجد :

أ- $f_{0.05,4,60}$ ب- $f_{0.025,8,10}$ جـ- $f_{0.95,4,6}$ د- $f_{0.99,20,12}$.

الحل :

أ - من جدول (7) و عندما $\alpha = .05$ و $m = 4$ و $n = 6$ نجد أن : $f_{0.05,4,6} = 4.53$.

ب - من جدول (7) و عندما $\alpha = 0.025$ و $m = 8$ و $n = 10$ نجد أن : $f_{0.025,8,10} = 3.85$

جـ - حيث أن جدول (7) يعطى القيم التي تقع على يمين التوزيع و عليه من العلاقة :

$$f_{1-\alpha,m,n} = \frac{1}{f_{\alpha,n,m}}$$

نجد أن :

$$f_{0.95,4,6} = \frac{1}{f_{0.05,6,4}} = \frac{1}{6.16} = 0.162 \text{ .}$$

د - بالمثل كما في (جـ) نجد أن : $f_{0.99,20,12} = \frac{1}{f_{0.01,12,20}} = \frac{1}{3.23} = 0.3096$

مثال (42) : إذا كان $F_{m,n}$ ترمز لمتغير عشوائي يتبع توزيع F بدرجات حرية تساوى m و n أوجد : أ - $P(F_{5,5} \le 5.05)$ ب - $P(F_{60,40} \ge 1.61)$.

الحل :

أ - من جدول (6) نجد أن : $P(F_{5,5} \le 5.05) = 0.05$.

ب - من جدول (6) : $P(F_{20,40} \ge 1.61) = 1 - P(F_{20,40} \le 2.01) = 1 - 0.10 = 0.90$

تمرينات Exercises

1 - إذا كان من بين 120 متقدم لوظيفة معينة 80 منهم مؤهلين لهذه الوظيفة فإذا تم اختيار 5 منهم بطريقة عشوائية أوجد الاحتمالات التالية باستخدام التوزيع فوق الهندسي والتوزيع ذات الحدين :

أ - احتمال أن 2 منهم يكونا مؤهلين .

ب - احتمال أن لا أحد منهم مؤهل .

د - احتمال على الأقل ثلاثة منهم مؤهلين .

2 - إذا كان احتمال نجاح متدرب في امتحان قيادة السيارة هو 0.40 ما احتمال :

أ - أن المتدرب العاشر الذي نجح في الامتحان يكون ثالث واحد .

ب - أن المتدرب الثاني عشر الذي نجح في الامتحان يكون خامس واحد .

د - نجاح متدرب في المرة الرابعة .

3 - أوجد التوزيع المنتظم لعينات عشوائية حجمها 4 تم اختيارها من 6 طلاب .

4 - إذا علمت أن احتمال شفاء مريض بالقلب هو 0.8 ما احتمال :

أ - شفاء 5 من بين 7 أشخاص .

ب - شفاء على الأكثر ثلاثة من بين 7 .

د - احتمال عدم شفاء أي من السبعة أشخاص .

5 - إذا كانت الاحتمالات 0.2 ، 0.3 ، 0.5 ترمز على التوالي لاحتمال وصول طالب للكلية عن طريق السيارة ، سير على الأقدام ، باستخدام الدراجة ما احتمال أنه من بين 9 طلاب تم اختيارهم بطريقة عشوائية أن 3 منهم قادمين عن طريق السيارة ، 4 منهم قادمين عن طريق السير على الأقدام ، 2 منهم قادمين عن طريقة الدراجة .

6 - أظهرت الدراسة التي قامت بها الشركة العامة للبريد أن من بين 5000 هاتف يوجد 4000 عليها ديون مستحقة للشركة فإذا تم اختيار عشرة هواتف بطريقة عشوائية ما احتمال ثلاثة منها ليس عليها ديون (استخدم تقريب ذات الحدين للتوزيع فوق الهندسي) .

7 - إذا ألقيت عملة نقدية معدنية بشكل متكرر ما احتمال :

أ - الحصول على ثالث صور في الرمية السابعة .

ب - الحصول على أول صورة في الرمية الرابعة .

8 - إذا علمت أن متوسط عدد الحوادث بتقاطع ما يساوي 4 في كل شهر ما احتمال وقوع :

أ - 5 حوادث خلال شهر ما .

ب - أقل من 3 حوادث خلال شهر ما .

جـ - على الأقل 4 حوادث خلال شهر ما .

د - عدم وقوع حادث خلال شهر ما .

9 - إذا علمت أن سكرتيرة ترتكب في المتوسط 3 أخطأ بكل صفحة عند الطباعة ما احتمال أنها ترتكب في صفحة قادمة سوف تطبعها :

أ - 3 أخطأ أو أكثر .

ب - ولا خطأ .

جـ - على الأقل 5 أخطأ .

10 - بفرض أن X متغير عشوائي يتبع التوزيع الهندسي باحتمال نجاح في أي محاولة يساوي 0.1 أوجد :

أ - $P(X \geq 2)$.

ب - $P(X < 4)$.

جـ - $P(X > 4 | X > 2)$ ثم قارن الناتج مع $P(X > 2)$.

11 - إذا علمت أن 10 ٪ من المحركات الكهربائية المنتجة من قبل أحد المصانع بهـا خلل فـإذا علمت أن المحركات يتم اختيارها بطريقة عشوائية واحدة في كل مرة وذلك لأجـل اختيـاره أوجـد احتمال :

أ - أن أول محرك يتم اختياره و ليس به خلل يكون في المحاولة الثانية .

ب - ثالث محرك يتم اختياره و ليس به خلل في المحاولة الخامسة .

جـ - ثالث محرك يتم اختياره و ليس به خلل يكون في المحاولة الخامسة أو قبلها .

12 - إذا علمت أن معدل عدد المكالمات الهاتفية التي تسـتقبلها إدارة الكليـة هـو 5 مكالمـات فـي الساعة أوجد احتمال :

أ - استقبال على الأقل مكالمتين خلال ساعتين معينتين .

ب - استقبال على الأكثر 4 مكالمات خلال ساعة معينة .

جـ - عدم استقبال أي مكالمة خلال ساعة معينة .

د - استقبال على الأقل مكالمتين خلال ساعتين .

13 - يحتوي صندوق على 4 كرات بيضاء و 3 كرات حمراء سحبت كرتان بطريقة عشوائية وبدون إعادة أوجد احتمال :

أ - كرة واحدة بيضاء يتم سحبها .

ب - على الأقل كرة بيضاء يتم سحبها .

جـ - سحب كرتان بيضاوان بشرط أن كرة واحدة بيضاء على الأقل تم سحبها .

د - ثاني كرة يتم سحبها تكون بيضاء .

14 - تشير الإحصاءات المأخوذة عن مكتب إطفاء الحرائـق أن 73 ٪ من الحرائق تحدث فـي المنازل و 20 ٪ تحدث في السيارات و 7 ٪ تحدث في الغابات ، فإذا أستقبل المكتب 8 مكالمـات مستقلة عن حرائق خلال يوم معين فما احتمال أن 4 منها في المنازل و 3 منها في السيارات و واحدة في الغابات .

15 - إذا كان من بين عدد كبير من المتقدمين لوظيفة معينة 60 ٪ منهم يحملون شهادة ثانوية ، 30 ٪ منهم يحملون مؤهل جامعي و 10 ٪ منهم يحملون دبلوم متوسط فإذا تم اختيار 10 أشخاص من بين المتقدمين أوجد احتمال :

أ - 3 منهم يحملون مؤهل جامعي و 5 شهادة ثانوية و 2 دبلوم متوسط .

ب - على الأقل 6 منهم يحملون مؤهل جامعي .

16 - إذا علمت أنه بإمكان الزبائن الخروج من أحد المصارف عن 3 أبواب هي A ، B ، C و على افتراض أن أي زبون سوف يختار الخروج من أي باب باحتمال متساوي ما احتمال أنه من بين 6 زبائن :

أ - 2 يخرجان من A و 3 يخرجون من B و واحد من C .

ب - جميع الزبائن يخرجون من نفس الباب .

17 - إذا علمت أن شدة الهزة الأرضية التي ضربت أحد المناطق يمكن تمثيلها بالتوزيع الأسي بمتوسط يساوي 2.4 على مقياس رختر ما احتمال أن الهزة الأرضية القادمة التي تضرب نفس المنطقة يكون :

أ - أكبر من 3 على مقياس رختر ؟

ب - ما بين 2 و 3 على مقياس رختر ؟

جـ - أقل من 4 على مقياس رختر ؟

18 - أحد العاملين على خازنات المياه لاحظ أن الكمية المطلوب من الماء في ساعة معينة من اليوم يمكن تمثيلها بمتغير عشوائي يتبع التوزيع الأسي بمتوسط يساوي 100 قدم مكعب بكل ثانية ما احتمال :

أ - أن كمية المياه المطلوبة ستزيد عن 200 قدم مكعب خلال يوم ما .

ب - أن كمية المياه المطلوبة ستكون ما بين 150 ، 250 قدم مكعب خلال يوم ما .

19 - إذا علمت أن الزمن (X) المطلوب للقيام بإصلاح جهاز مرئي يتوزع وفق التوزيع الأسي بمتوسط يساوي 5 ساعات و أن تكاليف إصلاح جهاز مرئي يمكن تمثيلها بالعلاقة الآتية $C = 20 + 8X + X^2$. فأوجد : $E(C)$ و $V(C)$.

20 - إذا علمت أن احمالي كمية الأمطار التي تسقط على مدينة قصر الأخيار خلال الأسبوع التوزيع الأسي بمتوسط يساوى 20 ملم . أوجد احتمال أن :

أ - احمالي كمية الأمطار التي تسقط خلال اسبوع على هذه المدينة أكبر من 22 ملم .

ب - احمالي كمية الأمطار التي تسقط خلال أسبوع على هذه المدينة أقل من 20 ملم .

جـ - احمالي كمية الأمطار التي تسقط خلال أسبوع على هذه المدينة ما بين 18 و 23 ملم .

21 - إذا كان $X \sim N(3 . 16)$ أوجد :

أ - $P(4 \le X \le 8)$

ب - $P(-2 \le X \le 1)$

جـ - $P(0 \le X \le 5)$.

22 - إذا كان $Z \sim N(0 . 1)$ أوجد :

أ - $P(0 \le Z \le 2)$ و $P(1.25, Z \le 2.75)$ و $P(-1.65 \le Z \le 0.70)$.

ب - أوجد قيمة الثابت a بحيث :

1) $P(0 \le Z \le a) = 0.4147$.

2) $P(Z > a) = 0.05$.

3) $P(|Z| < a) = 0.95$.

23 - إذا كان $X \sim N(6.25)$ فأوجد : $P(0 \le X \le 8)$ ، $P(6 \le X \le 12)$ ، $P(X > 21)$ ، $P(|X-6| \le 5)$ ، $P(|X-6| \le 10)$ ، $P(|X-6| \le 15)$.

24 - بفرض أن X متغير عشوائي بدالة مولدة العزوم كما يلي :

$m(t) = e^{166t + 200t^2}$ فأوجد :

أ - $E(X)$ و $V(X)$.

ب - $P(170 < X < 200)$.

جـ - $P(148 \le X \le 172)$.

25 - من جدول t و بدرجات حرية تساوى 7 أوجد :

أ - $P(T \leq 1.415)$.

ب - $P(T \leq -1.415)$.

جـ - $P(-1.895 < T < 1.415)$.

26 - من جدول t أوجد :

أ - $P(-1.812 \leq T \leq 1.812)$ عندما $n = 10$

ب - الثابت a حيث $P(|T| < a) = 0.90$ عندما $n = 14$

27 - من جدول t أوجد :

أ - $P(T \geq 2.228)$ عندما $n = 10$.

ب - $P(|T| \geq 2.228)$ عندما $n = 10$.

جـ - $P(1.330 \leq T \leq 2.552)$ عندما $n = 18$.

د - $P(T \leq 2.228)$ عندما $n = 10$.

هـ - $P(-1.753 \leq T \leq 2.602)$ عندما $n = 15$.

28 - من جدول F أوجد :

أ - $P(F \geq 3.50)$ عندما $m = 7$ و $n = 8$

ب - $P(F \geq 14.66)$ عندما $m = 9$ و $n = 4$

جـ - $P(F \geq 0.244)$ عندما $m = 6$ و $n = 9$

د - $P(F \leq 0.40)$ عندما $m = 2$ و $n = 7$

29 - من جدول F أوجد a و b بحيث :

أ - $P(a \leq F \leq b) = 0.90$ عندما $m = 8$ و $n = 6$

ب - $P(a \leq F \leq b) = 0.98$ عندما $m = 8$ و $n = 6$

(ألحظ أن $P(F \leq c) = \frac{\alpha}{2}$ و $P(F \leq d) = 1 - \frac{\alpha}{2}$ و بالتالي فإن $P(c \leq F \leq d) = 1 - \alpha$).

30 - من جدول F أوجد كل من :

أ - $F_{4,6,0.05}$ ب - $F_{4,6,0.95}$

ج - $F_{12,12} = 0.005$ د - $F_{12,12,0.995}$

31 - من جدول مربع كاي أوجد كلا من :

أ - $\chi^2_{0.05,10}$ ب - $\chi^2_{0.01,5}$ جـ - $\chi^2_{0.30,20}$

د - $\chi^2_{0.70,20}$ هـ - $\chi^2_{0.95,30}$ و - $\chi^2_{0.99,30}$

32 - من جدول مربع كاي أوجد كلا من :

أ - $P\left(\chi^2_{(13)} \geq 8.634\right)$ ب - $P\left(\chi^2_{(12)} \geq 14.011\right)$

جـ - $P\left(8.542 \leq \chi^2_{(15)} \leq 27.488\right)$ د - $P\left(\chi^2_{(23)} \leq 32.007\right)$

33 - إذا علمت أن اجمالي كمية الأمطار التي سقطت لمدة أربعة أسابيع على مدينة قصر الأخيار يمكن تمثيلها بتوزيع جاما بمعلمتين $\alpha = 1.6$ ، $\beta = 2.0$.

أ - أوجد القيمة المتوقعة و التباين لأجمالي كمية الأمطار التي سقطت لمدة أربعة أسابيع على هذه المدينة .

ب - أوجد احتمال أن يكون اجمالي كمية الأمطار التي سقطت لمدة أربع أسابيع على هذه المدينة أقل من 8.26

جـ - إذا كانت $C = 3 + 5X$ أوجد $E(C)$ و $V(X)$.

34 - إذا كان الأس في التوزيع الطبيعي الثنائي كما يلي :

$$-\frac{1}{120}\left[(x+2)^2 - 2.8(x+2)(y-1) + 4(y-1)^2\right]$$

أوجد : أ - μ_1 , μ_2 , σ_1 , σ_2 , ρ .

ب - $\mu_{y/x}$, $\sigma^2_{y/x}$.

35 - إذا كان الأس في التوزيع الطبيعي الثنائي كالآتي :

$$-\frac{1}{54}\left[x^2+4y^2+2xy+2x+8y+4\right]$$

أوجد σ_1 ، σ_2 ، ρ علما بأن $\mu_1=0$ ، $\mu_2=-1$

36 - إذا كان X ، Y لهما توزيع طبيعي ثنائي وكان

$U=X+Y$ ، $W=X-Y$ أوجد $\rho_{U,W}$

37 - إذا علمت أن 80 ٪ من العمليات الجراحية التي تجرى بأحد المستشفيات ناجحة ، فإذا تم اختيار مريضان ستجرى لهما عمليات جراحة في الأسبوع القادم ، وكان المتغير العشوائي X يمثل عدد العمليات الجراحية الناجحة فأوجد :

i - التوزيع الاحتمالي للمتغير X .

ii - دالة التوزيع المتجمع التراكمي .

iii - القيمة المتوقعة و التباين لهذا التوزيع .

38 - إذا كانت نسبة المصابين بعمى الألوان في مجتمع ما هي 20 ٪ فأوجد :

i - التوزيع الاحتمالي لعدد المصابين بعمى الألوان في أسرة مكونة من 4 أفراد من هذا المجتمع

ii - ما هو متوسط و تباين هذا التوزيع الاحتمالي .

39 - في أحد الدراسات وجد أن 15٪ من الناس يستخدمون اليد اليسرى ، في عينة حجمها عشرة أشخاص سحبت من هذا المجتمع ، أوجد :

i - احتمال شخصان يستخدمون اليد اليسرى .

ii - احتمال على الأقل اثنان يستخدمون اليد اليسرى.

iii - احتمال على الأكثر ثلاثة يستخدمون اليد اليسر .

vi - أوجد التوقع و التباين في الحالة رقم (iii) .

40 - أوجد قيمة p . n لمتغير ذات الحدين إذا كان $\mu=5$ و $\sigma^2=\frac{15}{4}$.

41 - إذا كان متوسط عدد الأيام التي يقفل فيها مستوصف صحي أبوابه بسبب الصيانة هو أربعة أيام في السنة فما هو احتمال أن المستوصف سيقفل أبوابه ستة أيام في السنة القادمة بسبب الصيانة .

42 - إذا كان معدل عدد الحوادث التي يستقبلها قسم الحوادث بمستشفى طرابلس المركزي هو أربعة حوادث في اليوم الواحد . فما احتمال :

i - حدوث ثلاثة حوادث أو أقل في يوم معين .

ii - أن يزيد عدد الحوادث عن 2 خلال يومين وأوجد التوقع و التباين في هذه الحالة

43 - إذا كان معروف من خلال السنوات الماضية أن متوسط عدد المصابين بمرض معين خلال السنة هو 15 وكان عدد الوافيات من هذا المرض يتبع توزيع بواسون فما احتمال وفاة :

i - سبعة أشخاص خلال السنة .

ii - على الأقل خمسة أشخاص في السنة .

iii - على الأكثر خمسة أشخاص خلال السنة .

44 - إذا كانت نسبة الإصابة الأمراض الخطيرة في بلد ما هي واحد في الألف ما احتمال وجود على الأكثر مصاب بهذا المرض في منطقة سكنها 3000 نسمة أستخدم :

أ - توزيع ذات الحدين .

ب - توزيع بواسون كتقريب لذات الحدين و قارن بينهما .

45 - إذا كان معدل إجراء عمليات مستعصية في أحد المستشفيات هو خمسة عمليات في الشهر الواحد ، فأوجد احتمال أن يكون في الشهر القادم :

i - ستة عمليات بالضبط .

ii - ثلاثة عمليات فأكثر .

iii - عدم وقوع أي عملية جراحية .

iv - أكثر من اثنان و أقل من خمسة .

v - أوجد متوسط ، تباين هذا التوزيع .

احسب الاحتمالين الآتيين : إذا علمت أن احتمال إنجاب ولد يساوي 0.5

46 - أ - أن تكون لأسرة ما ولدان على الأقل من بين أربعة أطفال .

ب - أن يكون لأسرة ما ولد واحد و بنت واحدة على الأقل من بين خمسة أطفال .

47 - احسب التباين و الانحراف المعياري لتوزيع ذي حدين متوسطه 2.5 إذا كانت $n = 10$

48 - إذا كان المتغير العشوائي X يتوزع وفق التوزيع ذا الحدين بمتوسط يساوي 2 و تباين يساوي $\frac{3}{4}$. أكتب قيم هذا المتغير و احتمال الحصول على كل منها .

49 - إذا علمت أن 80٪ من الأشخاص المتبرعين بالدم إلى أحد المستشفيات يحملون فصيلة الدم A . فإذا تم اختيار 5 أشخاص عشوائياً من المتبرعين بالدم خلال يوم معين فأوجد احتمال

أ - أن يكون أحدهم على الأقل يحمل فصيلة الدم A .

ب - أن يكون أربعة منهم على الأكثر يحملون فصيلة الدم A .

50 - تقدم أربعة أشخاص لأحد الامتحانات ، فإذا علمت أن احتمال نجاح الشخص في هذا الامتحان يساوي 0.75 .

أ - أكتب التوزيع الاحتمالي لعدد الناجحين من بين الأشخاص الأربعة .

ب - أعرض التوزيع الناتج بيانياً .

جـ - أحسب المتوسط الحسابي و التباين .

د - أحسب احتمال نجاح شخصين على الأقل من بين المتقدمين الأربعة .

51 - إذا كان $X \sim G(P = .3)$ فأوجد :

أ - $P(X \geq 3)$.

ب - $P(X > 6 / X > 3)$ ثم قارن ذلك مع $P(X > 3)$.

52 - إذا كان 10 ٪ من الإطارات الموجودة بمخازن شركة الإطارات غير صالحة للاستعمال فإذا كان المطلوب اختيار أربعة إطارات من هذه المخازن لتركيبها بإحدى السيارات فأوجد

أ - احتمال اختيار 6 إطارات للحصول على أربعة صالحة للاستعمال .

ب - القيمة المتوقعة و التباين لعدد الاختيارات المطلوبة للحصول على أربعة إطارات صالحة للاستعمال .

53 - إذا كان $X \sim NB(P=.6)X$ فأوجد $P(X \geq 3)$ عندما : أ - $r=5$ ب - $r=8$.

54 - إذا علمت أن 30٪ من المتبرعين بالدم في أي يوم لأحد المستشفيات يحملون فصيلة الدم B ، فإذا تم اختيار يوم معين بشكل عشوائي فأوجد احتمال أن يكون :

أ - أول متبرع في ذلك اليوم يحمل الفصيلة B سيكون رابع متبرع .

ب - ثاني متبرع في ذلك اليوم يحمل الفصيلة B سيكون رابع متبرع .

55 - يوجد بإحدى الشركات بمدينة طرابلس مائة موظف ، 60 منهم يقطنون خارج مدينة طرابلس . اخذت عينة حجمها 5 أشخاص عشوائياً من داخل الشركة فما احتمال وجود :

أ - شخصين أو كثر يقطنون خارج طرابلس .

ب - شخص واحد يقطن داخل مدينة طرابلس .

جـ - على الأكثر شخصين يقطنون خارج مدينة طرابلس .

56 - إذا كان معدل وصول الزبائن إلى المتاجر هو 6 أشخاص كل ساعة ، فأوجد احتمال :

أ - وصول 4 أشخاص في ساعة معينة .

ب - وصول على الأقل شخصين في ساعة معينة .

جـ - وصول على الأكثر ثلاثة أشخاص في ساعة معينة .

د - وصول أكثر من خمسة أشخاص خلال ساعتين .

57 - إذا كان $X \sim P(\lambda = 3)$ فأوجد :

أ - $P(X=5)$ ب - $P(X \geq 5)$

جـ - $P(X<5)$ د - $P(X \geq 5 / X \geq 3)$.

58 - إذا كان متوسط عدد الصكوك التي يقوم أحد الموظفين بصرفها في أحد المصارف هو
صكوك كل أربع دقائق ، فما احتمال أن هذا الموظف سوف يقوم بصرف 6 صكوك خلال الأربع
دقائق القادمة ؟ وما احتمال عدم صرف أي صك ؟

59 - إذا كان عدد السيارات المارة على الطريق الساحلي عند نقطة معينة يتبع توزيع بواسون
بمعدل 5 سيارات كل دقيقة فأوجد احتمال :

أ - مرور 7 سيارات خلال دقيقتين .

ب - عدم مرور أي سيارة خلال دقيقتين .

جـ - مرور 6 سيارات خلال دقيقتين إذا علمت أنه قد مر أكثر من ثلاث سيارات .

د - مرور على الأقل 8 سيارات خلال ثلاث دقائق .

60 - إذا علمت أن نسبة الوحدات المعيبة في إنتاج سلعة ما يساوي 0.02 أحسب احتمال وجود
ثلاث وحدات معيبة في عينة مكونة من 200 وحدة . ثم احتمال وجود أكثر من ثلاث وحدات
معيبة في نفس العينة .

61 - يحتوي كتاب على 50000 كلمة من بينها 500 كلمة بها أخطأ مطبعية ، أحسب احتمال
عدم وجود أي خطأ في صفحة من هذا الكتاب بها 200 كلمة . ثم احسب احتمال وجود كلمة
واحدة في صفحة أخرى بها 250 كلمة .

62 - بفرض أن المتغير العشوائي X يتوزع وفق التوزيع الطبيعي بمتوسط يساوي واحد وتباين
يساوي 4 أوجد كل من :

أ - $P(|X| \le 3)$

ب - $P(2 < X < 6)$

جـ - $P(X \le 4)$

د - $P(|X| > 1.5)$

هـ - $P(1 \le -2X + 3 \le 8)$

و - $P(X < 0.5$ أو $X > 2)$

ى - قيمة a بحيث $P(|X - \mu| \le a) = 0.894$

63 - بفرض ان $Z \sim N(0,1)$ فأوجد قيمة k بحيث :

أ - $P(Z \le k) = 0.5$ ب - $P(Z \ge k) = 0.118$

جـ - $P(-k \le Z \le k) = 0.90$ د- $P(Z \le k) = 0.8749$.

64 - أوجد قيمة c بحيث كل دالة من الدوال الآتية تمثل دالة كثافة احتمال :

أ - $f_X(x) = c\,e^{-5x^2+9x-11}$ ، $x \in R$ ب - $f_X(x) = \frac{1}{\sqrt{\pi}} e^{-x^2+x-c}$, $x \in R$

65 - إذا علمت أن درجات الطلبة في مادة الإحصاء تتوزع وفق التوزيع الطبيعي بمتوسط يساوي 55 وانحراف معياري يساوي 8 ، وعلمت أن :

أ - أقل درجة للنجاح 47 فما هي نسبة الراسبين ؟

ب - 4.68 ٪ من الطلبة يحصلون على درجة ممتاز فما هي أقل درجة للحصول على تقدير ممتاز ؟

جـ - 80٪ من الدرجات العليا ناجحين فما هي درجة النجاح ؟



الفصل السادس

توزيعات المعاينة

Sampling Distributions

6-1 مقدمة Introduction

إن هذا الفصل يعد بمثابة حلقة وصل ما بين ما تعرضنا إليه سلفاً والاستنتاج الإحصائي الذي سوف نتعرض إليه في الفصول القادمة ، وبالتالي معرفة توزيعات المعاينة تعتبر مفتاح لفهم الاستنتاج الإحصائي .

عادةً ما يتم اختيار العينات وذلك لاكتشاف حقائق حول المجتمع الإحصائي قيد الدراسة ، هذه الحقائق يعبر عنها بدلالة أعداد تسمى معلمات (parameters) . هذه المعلمات عبارة عن كميات عددية تصف التوزيعات الاحتمالية (المجتمعات الإحصائية) وغالباً ما تكون مجهولة .

إن الاستنتاج الإحصائي غالباً ما يكون متعلق بدراسة هذه المعلمات سواء كان من حيث التقدير أو اختبارات الفروض الخاصة بها ، وللوصول لهذا الاستنتاج تستخدم إحصاءات وهي أعداد يمكن حسابها من بيانات العينة ، هذه الأعداد تعتمد على قيم العينة الداخلة في حساب تلك الإحصاءات ، وعليه فإن هذه الأعداد تتغير بتغير قيم العينة ، وبالتالي حكمها حكم المتغير العشوائي وحيث أنها كذلك يجب معرفة توزيع المعاينة لهذه الإحصاءات حتى نبرر استخدامها في الاستنتاج الإحصائي .

تعريف (1) : الإحصاءه (statistic) : هي أي دالة في مفردات العينة العشوائية ولا تعتمد على معلمة مجهولة .

فمثلاً : إذا كانت $X_1 , X_2 , X_3 ,\ldots, X_n$ تشكل عينة عشوائية عددها n فإن متوسط العينة $(\bar{X})$ وتباين العينة (S^2) والانحراف المعياري للعينة (S) ونسبة العينة $(\hat{p})$ جميعها تمثل إحصاءات . وحيث أن الإحصاءه دالة في بيانات من عينة عشوائية وعليه فهي متغير عشوائي وذلك لأنه وكما أشرنا أعلاه لبيانات مختلفة سيكون لهذه الإحصاءة قيم مختلفة وعلى العكس من ذلك فإن المعلمة ثابتة للمجتمع الواحد ولكنها تختلف من مجتمع إلى آخر .

تعريف (2) :توزيع المعاينة Sampling Distribution

إن التوزيع الاحتمالي لجميع القيم الممكنة لإحصاءة العينة يطلق عليه تسمية توزيع المعاينة .

إن هذا التوزيع الاحتمالي يخدم غرضين هما :

أ - يساعد في الإجابة على الاحتمالات المتعلقة بإحصاءة العينة .

ب- يعطي الجانب النظري الذي يبرر صحة تطبيق أساليب الاستنتاج الإحصائي .

وحيث أن أغلب الإحصاءات استخداماً في مجال الاستنتاج الإحصائي هي متوسط العينة وتباينها ونسبة عناصرها التي تحمل صفة معينة ، وبالتالي سوف نتناول في هذا الفصل توزيعات المعاينة لكل من متوسط العينة ونسبة عناصرها التي تحمل صفة معينة .

2-6 توزيع متوسط العينة Distribution of the Sample Mean

إن من أهم توزيعات المعاينة هو توزيع متوسط العينة ، وحيث أنه ما يهمنا هو معرفة صيغة هذا التوزيع ومتوسطه وتباينه ، وعليه سوف نوضح في هذا البند كيفية التعرف على صيغة توزيع المعاينة لمتوسط العينة حتى ولو بشكل تقريبي ، خاصة إذا كانت العينة مختارة من مجتمع توزيعه الاحتمالي غير معروف .

فإذا كان مجتمع (توزيع) المعاينة محدود وصغير فإن معرفة توزيع المعاينة للإحصاءة قد لا يكون صعباً ، ولكن إذا كان المجتمع الإحصائي كبيراً أو لانهائي فإنه ليس من الأمر السهل معرفة توزيع المعاينة ، ولكن ما نستطيع عمله في مثل هذه الحالة هو تقريب توزيع المعاينة لتلك الإحصاءة .إن توزيع المعاينة لأية إحصاءة يمكن الحصول عليه بالطرائق الرياضية ولكن هذا المدخل خارج نطاق هذا الكتاب .

تعريف (3) : إذا كانت $X_1, X_2, X_3, \ldots, X_n$ تمثل عينة عشوائية فإن متوسط العينة يرمز له بالرمز $\overline{X}$ ولتباينها بالرمز S^2 ، حيث :

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2 \quad \text{و} \quad \overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

فإذا كانت العينة العشوائية من مجتمع إحصائي (X) يخضع لتوزيع متوسطه μ وتباينه σ^2 فإن توزيع المعاينة لمتوسط العينة ($\overline{X}$) له الخصائص الآتية :-

أ - $\mu_{\overline{X}} = E(\overline{X}) = \mu$

ب - تباين متوسط العينة سيكون $\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n}$ إذا كان حجم المجتمع الإحصائي لانهائي أو كبيراً ، وسيكون $\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n}\left(\frac{N-n}{N-1}\right)$ إذا كان حجم المجتمع محدود . حيث N تمثل حجم المجتمع الإحصائي وان $\mu = \frac{\sum_{i=1}^{N} X_i}{N}$ و $\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(X_i - \mu)^2$.

لاحظ انه بالإمكان برهنة الخاصتين أعلاه باستخدام الأسلوب الرياضي ولكن سوف نكتفي بتوضيحهما من خلال مثال عددي .

مثال(1) : إذا كانت البيانات 3 ، 5 ، 1 ، 7 ، 9 تمثل مفردات مجتمع إحصائي وتم اختيار عينة عشوائية حجمها 2 فإنه يمكن أيجاد التوزيع الاحتمالي للإحصاءة $\overline{X}$ عندما تكون المعاينة مع الإعادة وبدون إعادة وذلك كما يلي :

لاحظ أولاً حيث أن هذا المجتمع به 5 مفردات فقط وعليه فإن

$$\mu = \frac{3+5+1+7+9}{5} = \frac{25}{5} = 5$$

وإن

$$\sigma^2 = \frac{1}{5}\left[(3-5)^2 + (5-5)^2 + (1-5)^2 \;\; (7-5)^2 + (9-5)^2\right]$$

$$= \frac{1}{5}\left[4+0+16+16+4\right] = \frac{40}{5} = 8$$

أ - عندما تكون المعاينة مع الإعادة فإن عدد العينات الممكنة الاختيار في هذه الحالة يساوي $5^2=25$ عينة وإن ومتوسط وتباين ونسبة القيم التي أقل من 5 بكل عينة ممكنة كما هو موضح في الجدول التالي :

جدول (1) : العينات الممكنة السحب عندما يكون الاختيار مع الإعادة .

الاختيار الأول	الاختيار الثاني: 1	3	5	7	9
1	$\bar{x}=1$ $s^2=0$ $\hat{p}=1$	$\bar{x}=2$ $s^2=2$ $\hat{p}=1$	$\bar{x}=3$ $s^2=8$ $\hat{p}=0.5$	$\bar{x}=4$ $s^2=18$ $\hat{p}=0.5$	$\bar{x}=5$ $s^2=32$ $\hat{p}=0.5$
3	$\bar{x}=2$ $s^2=2$ $\hat{p}=1$	$\bar{x}=3$ $s^2=0$ $\hat{p}=1$	$\bar{x}=4$ $s^2=2$ $\hat{p}=0.5$	$\bar{x}=5$ $s^2=8$ $\hat{p}=0.5$	$\bar{x}=6$ $s^2=18$ $\hat{p}=0.5$
5	$\bar{x}=3$ $s^2=8$ $\hat{p}=0.5$	$\bar{x}=4$ $s^2=2$ $\hat{p}=0.5$	$\bar{x}=5$ $s^2=0$ $\hat{p}=0$	$\bar{x}=6$ $s^2=2$ $\hat{p}=0$	$\bar{x}=7$ $s^2=8$ $\hat{p}=0$
7	$\bar{x}=4$ $s^2=18$ $\hat{p}=0.5$	$\bar{x}=5$ $s^2=8$ $\hat{p}=0.5$	$\bar{x}=6$ $s^2=2$ $\hat{p}=0$	$\bar{x}=7$ $s^2=0$ $\hat{p}=0$	$\bar{x}=8$ $s^2=2$ $\hat{p}=0$
9	$\bar{x}=5$ $s^2=32$ $\hat{p}=0.5$	$\bar{x}=6$ $s^2=8$ $\hat{p}=0.5$	$\bar{x}=7$ $s^2=8$ $\hat{p}=0$	$\bar{x}=8$ $s^2=2$ $\hat{p}=0$	$\bar{x}=9$ $s^2=0$ $\hat{p}=0$

ومن هذا الجدول نلحظ أن قيمة متوسط العينة تختلف من عينة إلى أخرى أي أن قيمته تعتمد على المفردات الداخلة في حسابه ، وبالتالي إنه يسلك سلوك متغير عشوائي وحيث أنه كذلك فسيكون لهذا المتغير توزيع احتمالي – وكذلك الأمر بالنسبة لتباين العينة ونسبة عناصرها التي تحمل صفة معينة – ، وحيث أنه لكل قيمة ممكنة من هذه القيم احتمال يساوى $\frac{1}{25}$ بأن تكون من ضمن قيم متوسط العينة وعليه فإن التوزيع الاحتمالي لمتوسط العينة $(\overline{X})$ يمكن وضعه في جدول وذلك كما يلي :

$\overline{X}=\overline{x}$	$P(\overline{X}=\overline{x})$	$\overline{x}P(\overline{X}=\overline{x})$	$(\overline{x}-\mu_{\overline{x}})^2P(\overline{X}=\overline{x})$
1	$\frac{1}{25}=0.04$	0.04	$(1-5)^2(0.04)=0.64$
2	$\frac{2}{25}=0.08$	0.16	$(2-5)^2(0.08)=0.72$
3	$\frac{3}{25}=0.12$	0.36	$(3-5)^2(0.12)=0.48$
4	$\frac{4}{25}=0.16$	0.64	$(4-5)^2(0.16)=0.16$
5	$\frac{5}{25}=0.20$	1.00	$(5-5)^2(0.2)=0$
6	$\frac{4}{25}=0.16$	0.96	$(6-5)^2(0.16)=0.16$
7	$\frac{3}{25}=0.12$	0.84	$(7-5)^2(0.12)=0.48$
8	$\frac{2}{25}=0.08$	0.64	$(8-5)^2(0.08)=0.72$
9	$\frac{1}{25}=0.04$	0.36	$(9-5)^2(0.04)=0.64$
	1.0	$\sum_{\overline{x}=1}^{9}\overline{x}P(\overline{X}=\overline{x})=5$	$\sum_{\overline{x}=1}^{9}(\overline{x}-\mu_{\overline{x}})^2P(\overline{X}=\overline{x})=4$

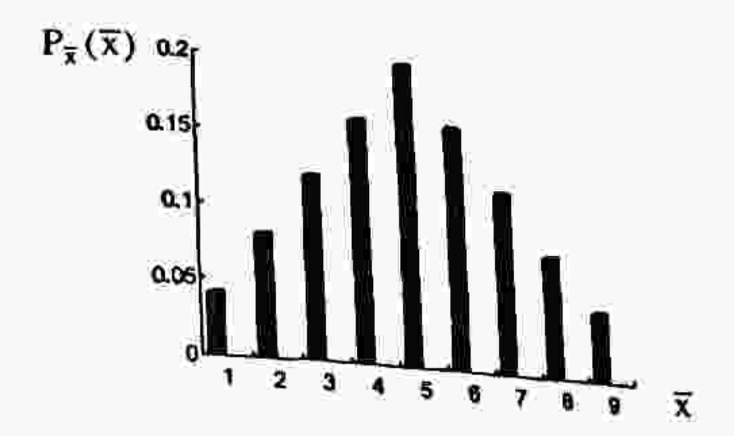
ومن هذا الجدول يتضح الآتي :

$$\mu_{\overline{x}} = E(\overline{X})=5=\mu$$

وإن

$$\sigma^2_{\overline{x}} = \sum_{x=1}^{9}(\overline{x}-\mu_{\overline{x}})^2 \; P(\overline{X}=\overline{x})=4$$

وإذا قمنا بقسمة تباين المجتمع الإحصائي على حجم العينة فإن الناتج سيكون مساوياً إلى $\sigma^2_{\overline{x}}$ ، أي أن $4=\frac{8}{2}=\frac{\sigma^2}{n}=\sigma^2_{\overline{x}}$ ومن هنا يتضح أن الخاصية (أ) والجزء الأول من الخاصية (ب) قد تحققا .ويمكن تمثيل التوزيع الاحتمالي لمتوسط العينة بيانياً كما يلي :



شكل (1) : التوزيع الاحتمالي لمتوسط العينة .

ومن هذا الشكل نلحظ أن التوزيع الاحتمالي متماثل .

ب - إذا كانت المعاينة بدون إعادة فإن عدد العينات الممكنة الاختيار في هذه الحالة يساوي $\binom{5}{2}=10$ عينات وإن ومتوسط ونسبة القيم التي أقل من 5 بكل عينة ممكنة كما هو موضح في الجدول التالي :

جدول (2) : العينات الممكنة السحب عندما يكون الاختيار بدون إعادة .

العينة	(1، 3)	(1،5)	(1،7)	(1،9)	(3،5)	(3،7)	(3،9)	(5،7)	(5،9)	(7،9)
$\bar{X}$	2	3	4	5	4	5	6	6	7	8
$\hat{p}$	1	0.5	0.5	0.5	0.5	0.5	0.5	0	0	0

وحيث أن احتمال اختيار أي عينة من العينات الممكنة أعلاه $=\frac{1}{10}$ وعليه فإن جدول التوزيع الاحتمالي لمتوسط العينة $(\bar{X})$ في هذه الحالة سيكون كالآتي :

$\overline{X}=\bar{x}$	$P(\overline{X}=\bar{x})$	$\bar{x}\,P(\overline{X}=\bar{x})$	$(\bar{x}-\mu_{\bar{x}})^2 P(\overline{X}=\bar{x})$
2	$\frac{1}{10}=0.1$	0.2	$(8-5)^2\,(0.1)=0.9$
3	$\frac{1}{10}=0.1$	0.3	$(3-5)^2\,(0.1)=0.4$
4	$\frac{2}{10}=0.2$	0.8	$(4-5)^2\,(0.2)=0.2$
5	$\frac{2}{10}=0.2$	1.0	$(5-5)^2\,(0.2)=0$
6	$\frac{2}{10}=0.2$	1.2	$(6-5)^2\,(0.2)=0.2$
7	$\frac{1}{10}=0.1$	0.7	$(7-5)^2\,(0.1)=0.4$
8	$\frac{1}{10}=0.1$	0.8	$(8-5)^2\,(0.1)=0.9$
	1.0	$\sum_{\bar{x}=2}^{8}\bar{x}\,P(\overline{X}=\bar{x})=5$	$\sum_{\bar{x}=2}^{8}(\bar{x}-\mu_{\bar{x}})^2 P(\overline{X}=\bar{x})=3$

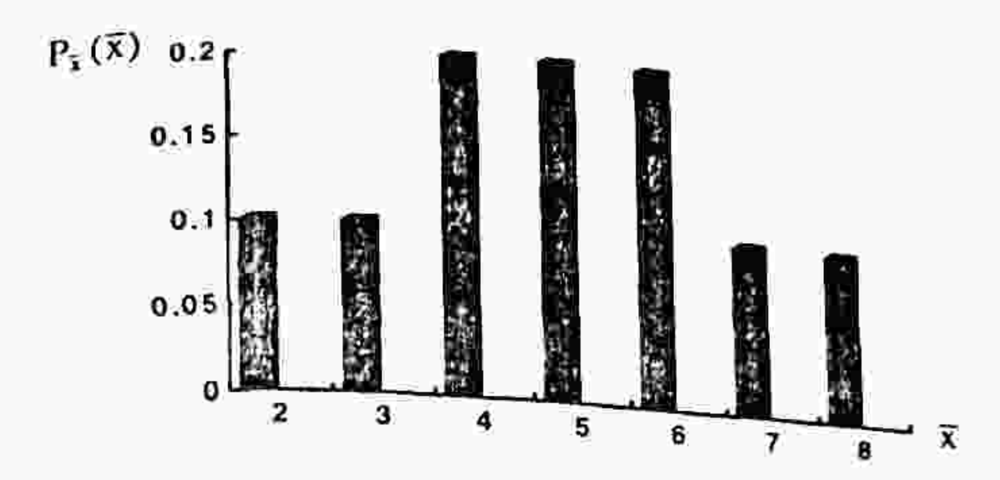
وبالتالي يتضح أن

$$\mu_{\bar{x}}=E(\overline{X})=\mu=5$$

وأن

$$\sigma_{\bar{x}}^2=\frac{\sigma^2}{n}\left(\frac{N-n}{N-1}\right)=\frac{8}{2}\left(\frac{5-2}{5-1}\right)=3$$

وبذلك تكون الخاصية (أ) والجزء الثاني من الخاصية (ب) قد تحققا ، ويمكن تمثيل التوزيع الاحتمالي بيانياً كما يلي :



شكل (2) : التوزيع الاحتمالي لمتوسط العينة .

ومن هذا الشكل نلحظ أن التوزيع الاحتمالي قريب من التماثل ، ومن خلال هذا المثال يتضح الآتي :

أ - إن أي إحصاءة مبنية على أساس عينة عشوائية هي متغير عشوائي ومصاحب لها توزيع احتمالي يسمى بتوزيع المعاينة .

ب - لاشتقاق توزيع المعاينة لأي إحصاءة نوجد جميع العينات الممكن اختيارها، ثم حساب قيمة الإحصاءة من كل عينة ، وحيث أن احتمال سحب كل عينة معروف فإنه يمكن تكوين جدول التوزيع الاحتمالي والذي يطلق عليه تسمية توزيع المعاينة للإحصاءة .

جـ - الحظ أن $\sigma_{\bar{X}}$ تقيس التشتت في توزيع المعاينة للإحصاءة $\bar{X}$ بينما σ تقيس التشتت في المجتمع الإحصائي (X) الذي سحبت منه جميع العينات الممكنة . ومن خلال هذا المثال نلاحظ أنه سواء كانت المعاينة مع الإعادة أو بدون إعادة فإن $\sigma_{\bar{X}}$ أصغر من σ وهذا يعني أن توزيع المعاينة لمتوسط العينة ($\bar{X}$) أقل تغايراً من مجتمع المعاينة .

د - في حالة المعاينة بدون إعادة إن المقدار $(\frac{N-n}{N-1})$ يطلق عليه تسمية معامل التصحيح المحدود ، وهذا المعامل يقترب من الواحد الصحيح عندما يكون حجم المجتمع الإحصائي كبيراً مقارنة بحجم العينة المختارة منه ، ولكن سواء كان المجتمع محدوداً أو غير محدود فإنه يفضل استخدام معامل التصحيح المحدود إذا كان حجم العينة أقل من 5 ٪ من حجم المجتمع الإحصائي .

هـ - إن التوزيع الاحتمالي في حالة ما يكون الاختيار بدون إعادة يكون أقل تغايراً منه في حالة الاختيار مع الإعادة ($\sigma_{\bar{X}}$ في حالة الاختيار مع الإعادة أكبر منه في حالة الاختيار بدون إعادة) وإذا نظرنا إلى الفرق ما بين متوسط العينة ومتوسط المجتمع الإحصائي نلاحظ أنه في حالة

الاختيار بدون إعادة أن أصغر المتوسطات يساوى 2 وبالتالي يكون الفرق ما بين هذا المتوسط و $\mu_{\bar{x}}$ مساوياً إلى 3 بينما في حالة الاختيار مع الإعادة إن أصغر المتوسطات يساوى 1 وبالتالي يكون الفرق ما بين هذا المتوسط و $\mu_{\bar{x}}$ مساوياً إلى 4 ومن هنا يتضح أن الاختيار بدون إعادة أفضل من الاختيار مع الإعادة .

و - هناك علاقة واضحة ما بين σ و $\sigma_{\bar{x}}$ فمن صيغتي $\sigma_{\bar{x}}$ (في الإعادة وبدون إعادة) نلاحظ أنه كلما كان المجتمع الإحصائي أكثر تغايراً (كلما كانت σ كبيرة) وكلما أدى ذلك لأن تكون $\sigma_{\bar{x}}$ كبيرة . وبعبارة أخرى كلما كان المجتمع أقل تغايراً كلما صغر حجم العينة المطلوب لتقدير متوسط المجتمع الإحصائي (μ) . أيضاً كلما كان حجم العينة كبيراً كلما كانت $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$ صغيرة وبالتالي يكون التوزيع الاحتمالي لمتوسط العينة أكثر تمركزاً أي أقل تشتتاً .

ز - إن توزيع المعاينة لمتوسط العينة $\bar{X}$ مبنى على جميع العينات العشوائية الممكنة ذات الحجم المعين التي يمكن اختيارها من المجتمع الإحصائي . ولكن في الواقع العملي سوف يتم اختيار عينة واحدة فقط وفي هذه الحالة هل توزيع $\bar{X}$ مفيد ؟ الإجابة نعم وذلك لأنه وكما رأينا عند اختيار عينة واحدة وحسب متوسطها فإن هذا المتوسط في الواقع هو واحد من بين جميع القيم الممكنة التي يمكن أن يأخذها متوسط العينة ، علاوة على ذلك ومن الجدول (1) إن هناك احتمال قدره 76 ٪ بأن متوسط العينة سوف لن يختلف عن متوسط المجتمع الإحصائي بأكثر من 2 ومن الجدول (2) . هذا الاحتمال يساوى 80 ٪ وعليه سنكون على ثقة إلى حد ما أنه أي متوسط عينة عشوائية بسيطة سوف لن يختلف عن متوسط المجتمع الإحصائي بشكل كبير ، وبهذه الكيفية يكون متوسط العينة يحتوى على جميع نتائج العينات الممكنة الاختيار ، وعلى أساس أي عينة عشوائية واحدة يتم اختيارها من المجتمع الإحصائي مدار البحث .

الحظ أنه في كثير من التطبيقات العملية لا يكون حجم المجتمع الإحصائي الذي سيتم منه اختيار العينة صغيراً ، ولكن سيكون كبير وبالتالي إيجاد جميع العينات الممكنة ذات الحجم المطلوب وتحديد التوزيع الاحتمالي لمتوسط العينة سوف لن يكون أمر سهلاً ،خاصة إذا كان التوزيع الاحتمالي توزيعاً غير طبيعي ، أما إذا كان مجتمع المعاينة يتبع التوزيع الطبيعي فإن توزيع المعاينة لمتوسط العينة يتبع التوزيع الطبيعي سواء كانت المعاينة مع الإعادة أو بدون إعادة .

والسؤال الذي يتبادر إلى الذهن الآن هو : هل هذا يعنى أنه إذا كان مجتمع المعاينة غير طبيعي لا يمكننا معرفة توزيع المعاينة لمتوسط العينة حتى ولو بشكل تقريبي من حيث مثلاً ، هل هو

متماثل أو قريب من التماثل ؟ إن الإجابة على هذا السؤال تكمن في النظرية التالية والتي تعد من أهم النظرية في مجال علم الإحصاء ، هذه النظرية يطلق عليها تسمية " نظرية النهاية المركزية . Central limit theorem

نظرية (1) :- إذا كانت $X_1 , X_2 , X_3 , ..., X_n$ عينة عشوائية من توزيع متوسطه μ وتباينه σ^2 محدود ، وكان حجم العينة (n) كبيراً فإن توزيع المعاينة للإحصاءة $\overline{X}$ يتوزع تقريباً وفق التوزيع الطبيعي بمتوسط يساوى μ $(\mu_{\overline{X}} = \mu)$ وتباين $\sigma^2_{\overline{X}} = \frac{\sigma}{n}$ إذا كانت المعاينة مع الإعادة و $\sigma^2_{\overline{X}} = (\frac{N-n}{N-1})\frac{\sigma^2}{n}$.

إن برهان هذه النظرية النهاية خارج نطاق هذا الكتاب ، وهي مهمة جداً من الناحتين النظرية والتطبيقية وذلك لسببين هما :

أولاً : إن الشرط في كون أن التباين يجب أن يكون محدود للتوزيع الذي سيتم منه اختيار العينة شرط محقق في معظم التوزيعات التي نواجهها في التطبيقات العملية .

ثانياً : إن حجم العينة (n) المطلوب لكي يكون التقريب جيد ليس كبيراً إلى حد ما ، وكقاعدة عامة إذا كانت $n \geq 30$ فإن توزيع المعاينة للإحصاءة $\overline{X}$ سيكون تقريبياً طبيعي إلى حد كبير . وفي الحقيقة إذا تم اختيار العينة من توزيع متصل ووحيد المنوال وملتويا قليلاً فإن توزيع $\overline{X}$ يمكن تقريبه بشكل جيد باستخدام التوزيع الطبيعي حتى إذا كان حجم العينة صغير إلى حد 5 أو 10 .

إن أهمية هذه النظرية ستكون واضحة إلينا في ما بعد عندما نعلم أن التوزيع الطبيعي أداة قوية في الاستدلال الإحصائي ، علاوة على ذلك نحن على ثقة على الأقل أن توزيع المعاينة لمتوسط العينة $(\overline{X})$ سيكون تقريباً توزيع طبيعي في أي حالة من الحالات الآتية :-

1 - عندما تكون العينة من مجتمع إحصائي توزيعه طبيعي .

2 - عندما تكون العينة من مجتمع إحصائي توزيعه غير طبيعي ولكن حجم العينة كبير .

3 - عندما تكون العينة من مجتمع إحصائي صيغة دالته غير معروفة ولكن حجم العينة كبير .

ملحوظ أخيرة تجدر الإشارة إليها هنا وهي أن معامل التصحيح المحدود ليس بالضرورة إذا كان $\frac{5}{n} \leq 0.05$ ، وفي معظم التطبيقات سوف لن يستخدم هذا المعامل وذلك بسبب تحقق هذا الشرط . وسوف نوضح النظرية السابقة من خلال دراسة الأمثلة التالية .

مثال (2) : مصنع لإنتاج النضائد السائلة يدعى أن متوسط أعمار النضائد التي ينتجهـا هـو 54 شهراً وبانحراف معياري يساوى 6 أشهر فإذا تم اختيار عينة عشوائية تتكون من 50 نضيدة مـن إنتاج هذا المصنع وذلك لغرض التحقق من صحة ادعاءه فأوجد احتمال أن يكـون متوسـط أعمـار عينة النضائد :

أ - أقل من 52 شهراً ؟ ب - أكثر من 55 شهراً ؟ جـ - بين 53 و 55 شهراً ؟

الحل :

حيث أن العينة التي تم اختيارها هي واحدة من بين جميـع العينـات الممكنـة ذات الحجـم 50 والتي يمكن سحبها من مجتمع النضـائد ، وبالتـالي فـإن المتوسـط الـذي سـتحصل عليـه مـن هـذه العينة هو أحد المتوسطات الداخلة في إيجاد توزيع المعاينة للإحصـاءة $\overline{X}$ الـذي يمكـن الحصـول عليه من مجتمع النضائد ، وسوف نفترض أن مجتمع النضائد مجتمع كبير مقارنـة بحجـم العينـة وعليه لا داعي لاستخدام معامل التصحيح .

إذن إذا كانت $\overline{X}$ تمثل متوسط أعمار النضائد للعينة المختارة عشوائياً فإنه وفقاً لنظرية النهاية المركزية يمكن القول بـأن $\overline{X}$ لهـا توزيـع طبيعـي تقريبيـاً بمتوسـط $\mu_{\overline{X}}=\mu=54$ وبـانحراف معياري $\sigma_{\overline{X}}=\frac{\sigma}{\sqrt{n}}=\frac{6}{\sqrt{50}}=0.85$. ولحساب المساحة تحت المنحنى الطبيعـي $N(\mu,\sigma^2)$ يجب معايرة المتغـير العشـوائي ليتحـول إلـى توزيـع طبيعـي معيـاري واسـتخدام جـداول التوزيـع الطبيعي المعياري لحساب الاحتمالات المطلوبة .

أ -

$$P(\overline{X}\le 52)=P(\frac{\overline{X}-\mu_{\overline{X}}}{\sigma_{\overline{X}}}\le\frac{52-54}{\sigma_{\overline{X}}})$$

$$=P(\frac{\overline{X}-54}{0.85}\le\frac{52-54}{0.85})$$

$$=P(Z\le -2.35)$$

ومن جدول التوزيع الطبيعي المعياري

$$= 0.5 - 0.4906 = 0.0094$$

ب -

$$P(\overline{X}\ge 55)=P(\frac{\overline{X}-\mu_{\overline{X}}}{\sigma_{\overline{X}}}\ge\frac{55-\mu_{\overline{X}}}{\sigma_{\overline{X}}})$$

$$= P(\frac{\overline{X}-54}{0.85} \geq \frac{55-54}{\sigma_{\overline{x}}})$$
$$= P(Z \geq 1.18)$$
$$= 0.5 - P(0 \leq Z \leq 1.18)$$
$$= 0.5 - 0.3810 = 0.119$$

جـ -
$$P(53 \leq \overline{X} \leq 55) = P(\frac{53-\mu_{\overline{x}}}{\sigma_{\overline{x}}} \leq \frac{\overline{X}-\mu_{\overline{x}}}{\sigma_{\overline{x}}} \leq \frac{55-\mu_{\overline{x}}}{\sigma_{\overline{x}}})$$
$$= P(\frac{53-54}{0.85} \leq \frac{\overline{X}-54}{0.85} \leq \frac{55-54}{0.85})$$
$$= P(-1.18 \leq Z \leq 1.18)$$
$$= 2P(0 \leq Z \leq 1.18)$$ من خاصية التماثل
$$= 2(0.3810) = 0.7620$$

مثال (3): بفرض أن X متغير عشوائي بدالة كتلة احتمال معرفة كما يلي :

$$p_X(x) = P(X=x) = \begin{cases} \frac{1}{4} & , x = 0,1,2,3 \\ 0 & \text{ow.} \end{cases}$$

فإذا تم اختيار عينة عشوائية حجمها 36 ومع الإعادة فما احتمال أن يكون متوسط هذه العينة أقل من 1.9 وأكبر من 1.5 إذا كان المتوسط مقاس لأقرب رقم عشري .

الحـــل :

حيث أن
$$\mu = E(X) = \sum_{x=0}^{3} x p(x) = \frac{1}{4}(0+1+2+3) = \frac{6}{4} = 1.5$$

و

$$\sigma^2 = E(X-\mu)^2 = \sum_{x=0}^{3}(x-1.5)^2 p(x)$$
$$= \frac{1}{4}\left[(0-1.5)^2 + (1-1.5)^2 + (2-1.5)^2 + (3-1.5)^2\right]$$
$$= \frac{1}{4}\left[2.25 + 0.25 + 0.25 + 2.25\right]\frac{5}{4} = 1.25$$

بالتوزيع المعاينـة للإحصـاءة $\overline{X}$ يمكن تقريبـه باستخدام التوزيـع الطبيعـي وذلـك وفقـاً لنظريـة

النهاية المركزية ، حيث $\mu_{\overline{x}}=\mu=1.5$ و $\sigma_{\overline{x}}=\frac{\sigma}{\sqrt{n}}=\frac{\sqrt{1.25}}{\sqrt{36}}=0.186$.

وحيث أن المقياس لأقرب رقم عشري وبالتالي فإن الاحتمال المطلوب سيكون كالآتي :

$$P(1.5<\overline{X}<1.9)\cong P\left(\frac{1.55-\mu_{\overline{x}}}{\sigma_{\overline{x}}}<\frac{\overline{X}-\mu_{\overline{x}}}{\sigma_{\overline{x}}}<\frac{1.85-\mu_{\overline{x}}}{\sigma_{\overline{x}}}\right)$$

$$=P\left(\frac{1.55-1.5}{0.186}<\frac{\overline{X}-1.5}{0.186}<\frac{1.85-1.5}{0.186}\right)$$

$$=P\left(\frac{1.55-1.5}{0.186}<Z<\frac{1.85-1.5}{0.186}\right)$$

$$=P(0.269<Z<1.88)$$

$$=P(0<Z<1.88)-P(0<Z<0.269)$$

$$=0.4699-0.1064=0.3635$$

مثال (4) : إذا كانت قيم الحامض البولي للأشخاص الطبيعيين تتوزع تقريباً وفق التوزيع الطبيعي بمتوسط يساوى 5.7 ملم٪ وانحراف معياري 1 ملم ٪ ، فإذا تم اختيار عينة عشوائية من 9 أشخاص فما احتمال أن يكون متوسط العينة :

أ- أكبر من 6 .   ب - أقل من 5.2 .   جـ - ما بين 5 و 6 .

الحل :

حيث أن المعاينة من مجتمع طبيعي وعليه فإن $\overline{X}$ التي تمثل متوسط قيم الحامض البولي بالعينة المختارة عشوائياً تتبع التوزيع الطبيعي بمتوسط يساوى 5.7 وتباين

$\sigma^2_{\overline{x}}=\frac{\sigma^2}{n}=\frac{(1)^2}{9}=\frac{1}{9}$ أي أن $\overline{X}\sim N(5.7,\frac{1}{9})$ وعليه فإن :

أ-

$$P(\overline{X}\geq 6)=P\left(\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}\geq\frac{6-\mu}{\sigma/\sqrt{n}}\right)$$

$$=P\left(\frac{\overline{X}-5.7}{1/\sqrt{9}}\geq\frac{6-5.7}{1/\sqrt{9}}\right)=P(Z\geq 0.90)$$

$$=0.5-P(0\leq Z\leq 0.90)$$

$$=0.5-0.3159=0.1814$$

ب -

$$P(\overline{X} \leq 5.2) = P\left(\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \leq \frac{5.2-\mu}{\sigma/\sqrt{n}}\right)$$

$$= P\left(\frac{\overline{X}-5.7}{1/\sqrt{9}} \leq \frac{5.2-5.7}{1/\sqrt{9}}\right) = P(Z \leq -1.50)$$

$$= 0.5 - P(0 \leq Z \leq 1.50)$$

$$= 0.5 - 0.4332 = 0.0668$$

(خاصية التماثل)

جـ -

$$P(5 \leq \overline{X} \leq 6) = P\left(\frac{5-\mu}{\sigma/\sqrt{n}} \leq \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \leq \frac{6-\mu}{\sigma/\sqrt{n}}\right)$$

$$= P\left(\frac{5-5.7}{1/\sqrt{9}} \leq \frac{\overline{X}-5.7}{1/\sqrt{9}} \leq \frac{6-5.7}{1/\sqrt{9}}\right)$$

$$= P(-2.10 \leq Z \leq 0) + P(0 \leq Z \leq 0.90)$$

$$= 0.4821 + 0.3159 = 0.7980$$

6 - 3 توزيع الفرق ما بين متوسطي عينتين

Distribution of the Difference Between Two Sample Means

في كثير من التطبيقات العملية تكون الأبحاث متعلقة بدراسة مجتمعين ، وعلى وجه التعيين تكون هذه الأبحاث متركزة على معرفة فيما إذا كان هناك فرق ما بين متوسطي مجتمعين ، أو معرفة مقدار الفرق بينهما ، فمثلاً إذا كان أحد المصانع يستورد المواد الخام من مصدرين مختلفين قد ترغب إدارة هذا المصنع في معرفة أي المصدرين في المتوسط يعطي مواد خام أكثر جودة، أو من الممكن أنه يوجد في أحد المصانع خطي إنتاج وترغب الإدارة في معرفة أي الخطين في المتوسط يعطي أكثر إنتاجاً ، أو مثلاً من الممكن وجود برنامجين مختلفين للتدريب على وظيفة معينة وترغب الجهة ذات الاختصاص معرفة أي البرنامجين في المتوسط يؤهل عاملين أكثر كفاءة ، فإذا تمكن المشرفين على هذه الأبحاث من التوصل إلى معرفة وجود فرق ما بين المتوسطين فإنه من المرغوب فيه بعد ذلك هو معرفة مقدار هذا الفرق ، وللإجابة على ذلك جرت العادة على اختيار عينتين أي اختيار عينة من كل مجتمع من المجتمعين مدار المقارنة ثم مقارنة الفرق ما بين متوسطي هاتين العينتين .

فإذا تم اختيار عينة عشوائية من مجتمع إحصائي يخضع لتوزيع متوسطه μ_1 وتباينه σ_1^2 واختيرت عينة عشوائية من مجتمع إحصائي أخر يخضع لتوزيع متوسطه μ_2 وتباين σ_2^2 وكان تباين المجتمعين معلوماً وكانت $\overline{X}$ ترمز لمتوسط العينة الأولى و $\overline{Y}$ ترمز لمتوسط العينة الثانية وكانت العينتين مستقلتين فإنه يمكن إيجاد التوزيع الاحتمالي للفرق ما بين متوسط العينتين $(\overline{X}-\overline{Y})$ وذلك من خلال اختيار جميع العينات الممكنة ذات الحجم m من المجتمع الأول وحساب متوسط العينة في كل مرة من مرات الاختيار، وبالمثل بالنسبة للمجتمع الثاني لعينات ذات الحجم n ، ثم أخذ جميع أزواج المتوسطات الممكنة واحدة من المجتمع الأول وواحدة من المجتمع الثاني وحساب الفرق في كل مرة ، ولكن وكما نعلم إن هذه الطريقة ستكون صعبة جداً حتى وإن كان حجم المجتمعين محدوداً ، وبالتالي لتفادى هذه الصعوبة سوف نفترض أن حجم العينتين كبيراً (ألحظ أنه إذا كان المجتمعين طبيعيين فإن هذا الشرط ليس بالضرورة لأن توزيع المعاينة للفرق ما بين متوسطي العينتين سيكون طبيعي أما إذا كانت الصيغ الدالية للمجتمعين غير معروفة فهذا الشرط ضروري) حتى يمكن تطبيق نظرية النهاية المركزية لأن هذه النظرية يمكن تعميمها لأكثر من عينة .

نظرية (2) :- إذا كانت $X_1, X_2, X_3, \ldots, X_m$ عينة عشوائية من مجتمع إحصائي يخضع لتوزيع (منفصل أو متصل) متوسطه μ_1 وتباينه σ_1^2 وكانت $Y_1, Y_2, Y_3, \ldots, Y_n$ عينة عشوائية من مجتمع إحصائي يخضع لتوزيع (منفصل أو متصل) متوسطه μ_2 وتباينه σ_2^2 وكانت العينتين مستقلتين و $\overline{X}$ ترمز لمتوسط العينة الأولى و $\overline{Y}$ ترمز لمتوسط العينة الثانية وكان حجم العينة كبيراً فإن توزيع المعاينة للفرق ما بين متوسطي العينتين يتوزع تقريباً وفق التوزيع الطبيعي بمتوسط $\mu_1-\mu_2$ وتباين $\sigma^2_{\overline{X}-\overline{Y}} = \frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}$.

إذن وفقاً هذه النظرية

$$(\overline{X}-\overline{Y}) \sim N(\mu_1-\mu_2 \ , \ \sigma^2_{\overline{X}-\overline{Y}})$$

وعليه فإن

$$Z = \frac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{\sqrt{\sigma^2_{\overline{X}-\overline{Y}}}} \sim N(0,1) \ .$$

لاحظ أنه إذا كانت $m \geq 30$ و $n \geq 30$ فإن التقريب الطبيعي لتوزيع $(\bar{X}-\bar{Y})$ سيكون جيد بصرف النظر عن شكل مجتمعي المعاينة .

مثال (5): إذا علمت أن الإنتاج السنوي لأحد مناجم الذهب يتوزع وفق التوزيع الطبيعي بمتوسط يساوي 150 طن وبانحراف معياري يساوي 20 طن بينما الإنتاج السنوي لمنجم آخر يتوزع وفق التوزيع الطبيعي بمتوسط يساوي 125 طن وبانحراف معياري يساوي 25 طن وتم اختيار عينة من إنتاج خمسة أشهر للمنجم الأول وعينة من إنتاج خمسة أشهر للمنجم الثاني ما احتمال:

أ - أن يكون متوسط عينة الإنتاج من المنجم الأول أصغر من أو يساوي متوسط عينة الإنتاج من المنجم الثاني .

ب - أن يكون الفرق ما بين متوسطي عينتي الإنتاج أكبر من أو يساوي 60 طن .

جـ - أن يكون الفرق ما بين متوسطي عينتي الإنتاج لا يقل عن 50 طن ولا يزيد عن 65 طن.

الحـــل :

بفرض أن $\bar{X}$ تمثل متوسط عينة الإنتاج من المنجم الأول ، $m = 5$.

$\bar{Y}$ تمثل متوسط عينة الإنتاج من المنجم الثاني ، $n = 5$.

إذن وفقاً للنظرية السابقة سيكون توزيع المعاينة للفرق ما بين متوسطي العينتين $(\bar{X}-\bar{Y})$ توزيعاً طبيعياً بمتوسط يساوي $\mu_1 - \mu_2 = 150 - 125 = 25$ وتباين يساوي $\sigma^2_{\bar{X}-\bar{Y}} = \frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n} = \frac{(20)^2}{5} + \frac{(25)^2}{5} = 205$. وعليه فإن

$$Z = \frac{(\bar{X}-\bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\sigma^2_{\bar{X}-\bar{Y}}}} = \frac{(\bar{X}-\bar{Y}) - 25}{\sqrt{205}} = \frac{(\bar{X}-\bar{Y}) - 25}{14.32} \sim N(0,1)$$

وبالتالي يمكن حساب الاحتمالات المطلوبة كما يلي :

أ -

$$P(\bar{X} \leq \bar{Y}) = P(\bar{X} - \bar{Y} \leq 0)$$

$$= P\left(\frac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{\sqrt{\sigma^2_{\overline{X}-\overline{Y}}}} \le \frac{0-(\mu_1-\mu_2)}{\sqrt{\sigma^2_{\overline{X}-\overline{Y}}}}\right)$$

$$= P\left(\frac{(\overline{X}-\overline{Y})-25}{14.32} \le \frac{0-25}{14.32}\right)$$

$$= P(Z \le -1.75)$$

$$= 0.5 - P(0 \le Z \le 1.75)$$

$$= 0.5 - 0.4599 = 0.0401$$

من خاصية التماثل

أي أن فرصة أن يكون متوسط عينة الإنتاج من المنتجم الأول أقل من أو يساوى متوسط عينة الإنتاج من المنتجم الثاني تساوى 4 ٪ .

ب-

$$P(\overline{X}-\overline{Y} \ge 60) = P\left(\frac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{\sqrt{\sigma^2_{\overline{X}-\overline{Y}}}} \ge \frac{60-(\mu_1-\mu_2)}{\sqrt{\sigma^2_{\overline{X}-\overline{Y}}}}\right)$$

$$= P\left(\frac{(\overline{X}-\overline{Y})-25}{14.32} \ge \frac{60-25}{14.32}\right)$$

$$= P(Z \ge 2.44)$$

$$= 0.5 - P(0 \le Z \le 2.44)$$

$$= 0.5 - 0.4927 = 0.0073$$

أي أن فرصة أن يكون الفرق ما بين متوسطي عينتي الإنتاج أكبر من أو يساوى 60 طن يساوى 0.1 ٪ تقريباً فقط .

جـ-

$$0 \le \overline{X}-\overline{Y} \le 65) = P\left(\frac{65-(\mu_1-\mu_2)}{\sqrt{\sigma^2_{\overline{X}-\overline{Y}}}} \le \frac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{\sqrt{\sigma^2_{\overline{X}-\overline{Y}}}} \le \frac{65-(\mu_1-\mu_2)}{\sqrt{\sigma^2_{\overline{X}-\overline{Y}}}}\right)$$

$$= P\left(\frac{50-25}{14.32} \le \frac{(\overline{X}-\overline{Y})-25}{14.32} \le \frac{65-25}{14.32}\right)$$

$$= P(1.75 \le Z \le 2.79)$$

$$= P(0 \le Z \le 2.79) - P(0 \le Z \le 1.75)$$

$$= 0.4974 - 0.4599$$

$$= 0.0375$$

أي أن فرصة أن يكون الفرق ما بين متوسطي عينتي الإنتاج أكبر من 50 طن وأقل من 65 طن هو 4 ٪ فقط .

مثال (6) : إذا كان المتوسط والانحراف المعياري لمستويات مصل الكولسترول للأشخاص الذين أعمارهم ما بين 25 سنة و 34 سنة يساوى 199 و 49 على التوالي ، بينما للأشخاص الذين أعمارهم ما بين 20 سنة و 24 سنة يساوى 180 و 43 على التوالي ، واختيرت عينتين عشوائيتين مستقلتين حجم كل منها 50 شخص من هذين المجتمعين ، فما احتمال أن يكون الفرق ما بين متوسط العينتين أكبر من 25 ؟

الحــل :

بفرض أن $\overline{X}$ تمثل متوسط مستويات مصل الكولسترول بالعينة المختارة من الأشخاص الذين أعمارهم ما بين 25 و 34 سنة . وبفرض أن $\overline{Y}$ تمثل مستويات مصل الكولسترول بالعينة المختارة من الأشخاص الذين أعمارهم ما بين 20 و 24 سنة .

حيث أن حجم العينة كبيراً وعليه من نظرية النهاية المركزية فإن توزيع المعاينة للفرق ما بين متوسطي العينتين $(\overline{X}-\overline{Y})$ سيكون توزيع طبيعي بمتوسط: $\mu_1-\mu_2=199-180=19$

وتباين $\sigma^2_{\overline{X}-\overline{Y}}=\frac{\sigma_1^2}{m}+\frac{\sigma_2^2}{n}=\frac{(49)^2}{50}+\frac{(43)^2}{50}=85$ ، وبالتالي فإن :

$$P((\overline{X}-\overline{Y})\geq 25)=P\left(\frac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{\sqrt{\sigma^2_{\overline{X}-\overline{Y}}}}\geq\frac{25-(\mu_1-\mu_2)}{\sqrt{\sigma^2_{\overline{X}-\overline{Y}}}}\right)$$

$$=P\left(\frac{(\overline{X}-\overline{Y})-19)}{\sqrt{85}}\geq\frac{25-19}{\sqrt{85}}\right)=P\left(Z\geq\frac{6}{9.2195}\right)$$
$$=P(Z\geq 0.65)=0.5-P(0\leq Z\leq 0.65)$$
$$=0.5-0.2422=0.2578$$

6 - 4 توزيع نسبة العينة Distribution of the sample proportion

لقد تعرضنا في البند السابق لتوزيعات المعاينة لإحصاءات يمكن قياسها وإعطائها قيم مختلفة ، واعتبارنا هذه القيم مستقلة عن بعضها البعض ، وكان الاهتمام متعلق بمعالجة مقادير هذه القيم بالذات . ولكن في بعض التطبيقات العملية لا يكون البحث منصباً على معالجة مقادير

تعتمد وإنما على معالجة عدد المفردات التي تدل عليها هذه القيم . وبعبارة أخرى فإن الدراسة هنا ستكون مترّكزة على معالجة عدد المفردات الواقعة في عدد من الفترات (أو الفئات) لظاهرة ما بغض النظر عن أطوال هذه الفترات (أو الفئات) أو مقاديرها ، ولكي نتعامل مع هذه البيانات إحصائياً يقترح الإحصائيين تحويل هذه البيانات إلى نسب أو معالجتها بطريقة معينة نستخلص منها قيمة يطلق عليها تسمية مربع كاي ، وإن هاتين الطريقتين مختلفتين حيث إن لكل منهما شروطها الخاصة التي تستعمل فيها ، وفي هذا البند سوف نتعرض للطريقة الأولى .

في بعض الأحيان يبدو من الضروري تقدير أو اتخاذ قرار يتعلق بنسبة عناصر أو مفردات المجتمع الإحصائي التي تحمل صفة معينة ، فمثلاً نسبة المصابين بمرض معين في مجتمع ما ، أو نسبة العاطلين عن العمل ، أو نسبة التلف من الإنتاج لسلعة ما ... الخ ، إن إحصاءة العينة التي عادةً ما تستخدم لإعطاء معلومات عن نسبة مفردات المجتمع الإحصائي التي تحمل صفة معينة يطلق عليه تسمية نسبة العينة (Sample proportion) وفي هذا البند سوف ندرس توزيع المعاينة لنسبة العينة على افتراض أن جميع بيانات العينة تم الحصول عليها من عينة عشوائية بسيطة .

إذا كانت X تمثل عدد مفردات أو عناصر المجتمع الإحصائي التي تمتلك الصفة مدار البحث و N تمثل عدد مفردات المجتمع الإحصائي ، فإن نسبة مفردات أو عناصر المجتمع الإحصائي التي تمتلك الصفة مدار البحث نرمز لها بالرمز p حيث $p=\dfrac{X}{N}$ وكانت $X_1, X_2, X_3, \ldots, X_n$ تشكل عينة عشوائية من المجتمع الإحصائي مدار البحث ، فإن نسبة عناصر أو مفردات العينة التي تمتلك الصفة مدار البحث يرمز لها بالرمز $\hat{p}$ حيث $\hat{p}=\dfrac{x}{n}$

وإن x تمثل عدد مفردات أو عناصر العينة التي تمتلك الصفة مدار البحث و n تمثل عدد مفردات العينة . ولهذه الإحصاءة الخصائص الآتية :

1- $\mu_{\hat{p}} = p$.

2- أ- إذا كانت المعاينة مع الإعادة فإن $\sigma^2_{\hat{p}} = \dfrac{p(1-p)}{n}$

حيث $\sigma^2 = p(1-p)$ تمثل تباين مجتمع الإحصاءة .

ب - إذا كانت المعاينة بدون إعادة فإن $\sigma^2_{\hat{p}} = \left(\dfrac{N-n}{N-1}\right)\left(\dfrac{p(1-p)}{n}\right)$ ولنتعرف على مدى صحة هذه الخصائص ، سوف نستخدم نفس البيانات التي سبق وأن تعرضنا إليها عند دراسة

توزيع المعاينة لمتوسط العينة . فإذا أردنا حساب نسبة المفردات التي أقل من 5 بالمجتمع الإحصائي مدار البحث فإن : نسبة مفردات المجتمع الإحصائي التي أقل من 5 تكون كالآتي : $p=\frac{2}{5}=0.4$. بينما تباين نسبة مفردات المجتمع الإحصائي التي أقل من 5 سيكون كالآتي :

$$\sigma^2=p(1-p)=(0.4)(1-0.4)=0.24$$

ولتوضيح السبب في كتابة التباين بهذه الصيغة الحظ أنه حيث أنه ما يهمنا هو إما أن يكون الرقم أقل من 5 أم لا ، وبالتالي إذا أعطينا العدد 0 إذا كانت المفردة أكبر من 5 والعدد 1 إذا كانت المفردة أقل من 5 ، فإن مفردات المجتمع الإحصائي 9 ، 7 ، 1 ، 5 ، 3 تتحول إلى : 0 ، 0 ، 1 ، 0 ، 1 . وبالتالي فإن متوسط عدد القيم التي أقل من 5 هو عبارة عن نسبة المفردات التي أقل من 5 ، أي أن

$$p=\frac{\sum_{i=1}^{N} X_i}{N} \quad \Rightarrow \quad \sum_{i=1}^{N} X_i = Np$$

ومن تعريف التباين نحن نعلم أن :

$$\sigma^2=\frac{\sum_{i=1}^{N}(X_i-\mu)^2}{N}=\frac{1}{N}\left[\sum_{i=1}^{N} X_i^2-\frac{(\sum_{i=1}^{N} X_i)^2}{N}\right]$$

وحيث أن X_i إما أن تكون صفر أو واحد وعليه فإن $\sum_{i=1}^{N} X_i=\sum_{i=1}^{N} X_i^2$ وبالتالي فإن

$$\sigma^2=\frac{1}{N}\left[Np-\frac{(Np)^2}{N}\right]=p(1-p)$$

إن توزيع المعاينة لنسبة مفردات العينة التي أقل من 5 في حالة ما تكون المعاينة مع الإعادة وبدون إعادة يكون كما يلي :

1 - عندما تكون المعاينة مع الإعادة :

لنفرض أنه تم اختيار عينة مكونة من مفردتين ومع الإعادة ونود إيجاد التوزيع الاحتمالي (توزيع المعاينة) لنسبة القيم التي أقل من 5 بهذه العينة $(\hat{p})$ فإنه وكما في حالة توزيع المعاينة لمتوسط العينة سوف نوجد أولاً كل العينات الممكنة ثم حساب النسبة في كل عينة ومنها يمكن

لإيجاد التوزيع المطلوب . حيث أن احتمال اختيار كل عينة هو $\frac{1}{25}$ وبحساب نسبة القيم التي أقل من 5 في كل عينة من هذه العينات كما هو موضح في الجدول (1) نجد أن التوزيع الاحتمالي لنسبة العينة $(\hat{p})$ يكون كما يلي :

$\hat{p}$	0	0.5	1	المجموع
$P(\hat{p})$	$\frac{9}{25}$	$\frac{12}{25}$	$\frac{4}{25}$	1

وعليه فإنه من تعريف التوقع الرياضي والتباين نجد أن

$$\mu_{\hat{p}} = E(\hat{p}) = \sum \hat{p}P(\hat{p})$$

$$=(0)(\frac{9}{25})+(0.5)(\frac{12}{25})=(1)(\frac{4}{25})=\frac{10}{25}=0.4$$

$$\sigma^2_{\hat{p}} = E(\hat{p}-\mu_{\hat{p}})^2 = \sum(\hat{p}-\mu_{\hat{p}})^2 P(\hat{p})$$

$$=(0-0.4)^2(\frac{9}{25})+(0.5-0.4)^2(\frac{12}{25})+(1-0.4)^2(\frac{4}{25})$$

$$=0.0576+0.0048+0.0576=0.12$$

وحيث أن $\sigma^2_{\hat{p}}=0.24$ وعليه بقسمة هذا المقدار على حجم العينة نجد أن :

$\frac{0.24}{2}=0.12$ اى ان $\sigma^2_{\hat{p}} = \frac{\sigma^2_p}{n} = \frac{p(1-p)}{n}$.

2 - عندما تكون المعاينة بدون إعادة :

لنفرض أنه تم اختيار عينة عشوائية تتكون من مفردتين وبدون إعادة فإن جميع العينات الممكنة ونسبة المفردات التي أقل من 5 موضحة في الجدول (2) ومن هذا الجدول نجد إن التوزيع الاحتمالي لنسبة العينة $(\hat{p})$ سيكون كالآتي :

$\hat{p}$	0	0.5	1	المجموع
$P(\hat{p})$	$\frac{3}{10}$	$\frac{6}{10}$	$\frac{1}{10}$	1

وعليه فإن

$$\mu_{\hat{p}} = E(\hat{p}) = \sum \hat{p}P(\hat{p})$$

$$=(0)(\frac{3}{10})+(0.5)(\frac{6}{10})+(1)(\frac{1}{10})=\frac{4}{10}=0.4$$

وإن

$$\sigma^2_{\hat{p}} = E(\hat{p}-\mu_{\hat{p}})^2 = \sum (\hat{p}-\mu_{\hat{p}})^2 P(\hat{p})$$

$$= (0-0.4)^2 (\frac{3}{10}) + (0.5-0.4)^2 (\frac{6}{10}) + (1-0.4)^2 (\frac{1}{10})$$

$$= 0.048 + 0.006 + 0.036 = 0.09$$

وحيث إن $\sigma^2_p = 0.24$ وإن

$$\sigma^2_p (\frac{N-n}{N-1}) = (\frac{0.24}{2})(\frac{5-2}{5-1}) = (0.24)(\frac{3}{4}) = 0.09$$

ومن هنا يتضح صحة الخصائص التي أشرنا إليها ، في الحقيقة أنه ليس بالضرورة إيجاد كل العينات الممكنة لاشتقاق توزيع المعاينة لنسبة العينة $(\hat{p})$ وذلك على خلاف توزيع المعاينة لمتوسط العينة $(\bar{X})$ والسبب في ذلك أن صيغة الدالة لتوزيع المعاينة لنسبة العينة $(\hat{p})$ تكون معروفة ، وهذه الصيغة تعتمد على ما إذا كان المجتمع محدود أو غير محدود . فإذا كان المجتمع محدود فإن توزيع المعاينة لنسبة العينة $(\hat{p})$ سيكون التوزيع فوق الهندسي الذي سبق وأن تعرضنا إليه ويمكن كتابته كما يلي :

$$P(\hat{p}) = P(\frac{X}{n}) = \begin{cases} \dfrac{\binom{NP}{x}\binom{N(1-p)}{n-x}}{\binom{N}{n}} & , x = 0,1,2,3,\ldots,n \\ 0 & , o.w \end{cases}$$

أما إذا كان المجتمع غير محدود (المعاينة مع الإعادة) فإن توزيع المعاينة لنسبة العينة $(\hat{p})$ سيكون توزيع ذي الحدين الذي سبق وأن تعرضنا إليه أيضاً وهو كما يلي :

$$P(\hat{p}) = P(\frac{X}{n}) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x} & , x = 0,1,2,3,\ldots,n \\ 0 & , ow \end{cases}$$

وبالنظر إلى التوزيع الاحتمالي لنسبة العينة $(\hat{p})$ سواء كانت المعاينة مع الإعادة أو بدون إعادة نلاحظ أن هذا التوزيع ملتوياً، وبصفة عامة ، إذا كانت P أقل من 0.5 فإن توزيع المعاينة للنسبة $(\hat{p})$ سيكون ملتوياً ناحية اليمين أما إذا كانت P أكبر من 0.5 فإن توزيع المعاينة للنسبة $(\hat{p})$ سيكون ملتوياً ناحية اليسار (في ما يخص التوزيع فوق الهندسي هذا صحيح إذا كان حجم العينة أقل من نصف حجم المجتمع الإحصائي وغالباً ما يكون كذلك في الحالات العملية)

ولكن بصفة عامة ، يمكن القول بأنه كلما زاد حجم العينة أقترب توزيـع المعاينـة للنسبة $(\hat{p})$ من التماثل بصرف النظر عن قيمة P وسيكون متماثل عندما تكون $P = 0.5$ وعليه إذا كان حجم العينة كبيراً إلى حدٍ ما فإن توزيع المعاينة لنسبة العينة $(\hat{p})$ سيكون توزيع طبيعي تقريباً.
إذن مما سبق وبتطبيق نظرية النهاية المركزية يمكن صياغة النظرية التالية وبدون برهان .

نظرية (3) :- إذا كانت $X_1, X_2, X_3, \dots, X_n$ عينة عشوائية من مجتمع إحصـائي بنسـبة P وكانت n كبيرة ، فإن توزيع المعاينة لنسبة العينـة $(\hat{p})$ يتوزيـع تقريبيـاً وفـق التوزيـع الطبيعـي بمتوسط $\mu_{\hat{p}} = p$ وتباين $\sigma^2_{\hat{p}} = \dfrac{p(1-p)}{n}$.

لاحظ أنه عندما تكون المعاينة من مجتمع محدود (أي أن المعاينة بدون إعادة) ونريـد تطبيـق هذه النظرية يجب أن يكـون حجـم المجتمـع أكـبر بكثـير مـن حجـم العينـة ، وإن القـاعدة العامـة لتطبيق هذه النظرية تقول لكي يكون التقريب باستخدام التوزيع الطبيعـي تقريـب مقبـول يجـب أن تكون قيمة كلا من np و $n(1-p)$ أكبر من 5 .

مثال (7) : إذا علمت أن 35 ٪ من الحوادث التي تحدث بمدينة مـا كـانت ناتجـة عـن السـرعة الفائقة لسائقي السيارات ، فإذا تم اختيار عينة عشوائية من 100 حادث مسجلة بأحد دوائر رجـال المرور فأوجد احتمال أن : أ - 45 ٪ منهم أو أكثر نتيجة للسرعة ؟ ب - ما بيـن 30 ٪ و 40 ٪ نتيجة للسرعة ؟

الحـل :

بفرض أن $(\hat{p})$ تمثل نسبة الحوادث التي حدثت نتيجة للسرعة بهذه العينـة ، وحيث أن حجـم العينة كبيراً وعليـه مـن النظريـة السـابقة يمكن الافتراض بـأن $(\hat{p})$ لهـا توزيـع طبيعـي تقريبـاً بمتوسط $\mu_{\hat{p}} = p = 0.35$ وتبـاين $\sigma^2_{\hat{p}} = \dfrac{p(1-p)}{n} = \dfrac{(0.35)(0.65)}{100} = 0.002275$ وعليـه فـإن

$$\sigma_{\hat{p}} = \sqrt{0.002275} = 0.0477$$

أ -

$$P(\hat{p} \geq 0.45) = P\left(\frac{\hat{p}-\mu_{\hat{p}}}{\sigma_{\hat{p}}} \geq \frac{0.45-\mu_{\hat{p}}}{\sigma_{\hat{p}}}\right) = P\left(\frac{\hat{p}-0.35}{\sigma_{\hat{p}}} \geq \frac{0.45-0.35}{\sigma_{\hat{p}}}\right)$$

$$=P(\frac{\hat{p}-0.35}{0.0477} \geq \frac{0.45-0.35}{0.0477})$$
$$=P(Z \geq 2.1)$$
$$=0.5-P(0 \leq Z \leq 2.1)$$
$$=0.5-0.4821=0.0179$$

ب-

$$P(0.30 \leq \hat{p} \leq 0.40)=P(\frac{0.30-\mu_{\hat{p}}}{\sigma_{\hat{p}}} \leq \frac{\hat{p}-\mu_{\hat{p}}}{\sigma_{\hat{p}}} \leq \frac{0.40-\mu_{\hat{p}}}{\sigma_{\hat{p}}})$$
$$=P(\frac{0.30-0.35}{0.0477} \leq \frac{\hat{p}-0.35}{0.0477} \leq \frac{0.40-0.35}{0.0477})$$
$$=P(-1.05 \leq Z \leq 1.05)$$
$$=2P(0 \leq Z \leq 1.05)$$
$$=2(0.3531)=0.7062$$

مثال (8) : إذا علمت أن 24 ٪ من المدخنين يفضلون تدخين سجائر من نوع خاص ، فإذا تم اختيار عينة عشوائية من 400 مدخن فأوجد احتمال :

أ- على الأكثر 30 ٪ منهم يفضلون هذا النوع من السجائر ؟

ب - على الأقل 28 ٪ منهم يفضلون هذا النوع من السجائر ؟

الحـــل :

بفرض أن ($\hat{p}$) تمثل نسبة المدخنين الذين يفضلون تدخين سجائر من نوع خاص بهذه العينة، وحيث أن حجم العينة كبير ، وعليه من النظرية السابقة فإن ($\hat{p}$) سيكون لها توزيع طبيعي تقريباً بمتوسط يساوى $\mu_{\hat{p}}=p=0.24$ وتباين يساوى

$$\sigma_{\hat{p}}^2=\frac{p(1-p)}{n}=\frac{(0.24)((0.76)}{400}=0.000456$$

وعليه فإن
$$\sigma_{\hat{p}}=\sqrt{0.000456}=0.0214$$

أ -

$$P(\hat{p} \leq 0.30)=P(\frac{\hat{p}-\mu_{\hat{p}}}{\sigma_{\hat{p}}} \leq \frac{0.30-\mu_{\hat{p}}}{\sigma_{\hat{p}}})$$
$$=P(\frac{\hat{p}-0.24}{0.0214} \leq \frac{0.30-0.24}{0.0214})$$
$$=P(Z \leq 2.8)=P(-\infty < Z \leq 0)+P(0 < Z \leq 2.8)$$

$$=0.5+0.4974=0.9974$$

ب -

$$P(\hat{p} \geq 0.28)=P(\frac{\hat{p}-\mu_{\hat{p}}}{\sigma_{\hat{p}}} \geq \frac{0.28-\mu_{\hat{p}}}{\sigma_{\hat{p}}})$$

$$=P(\frac{\hat{p}-0.24}{0.0214} \geq \frac{0.28-0.24}{0.0214})$$

$$=P(Z \geq 1.87)$$

$$=0.5-P(0 \leq Z \leq 1.87)$$

$$=0.5-0.4693=0.0307$$

مثال (9) : إذا علمت أن 35 ٪ من أفراد مجتمع ما يعانون من مرض واحد ومزمن أو أكثر ، فإذا تم اختيار عينة عشوائية من هذا المجتمع تتكون من 200 شخص ، فما احتمال أن 40 ٪ أو أكثر من أفراد هذه العينة يعانون على الأقل من مرض واحد ومزمن ؟

الحل :

بفرض أن $\hat{p}$ تمثل نسبة الأشخاص الذين يعانون على الأقل من مرض واحد ومزمن بهذه العينة ، حيث أنه من الواضح أن $nP > 5$ وعليه يمكن الافتراض بأن توزيع المعاينة لنسبة العينة ($\hat{p}$) يتبع التوزيع الطبيعي تقريبياً بمتوسط يساوي $\mu_{\hat{p}}=p=0.35$ وتباين يساوي

$$\sigma^2_{\hat{p}}=\frac{p(1-p)}{n}=\frac{(0.35)(1-0.35)}{200}=0.0011375$$

وعليه فإن

$$P(\hat{p} \geq 0.80)=P(\frac{\hat{p}-\mu_{\hat{p}}}{\sqrt{\sigma^2_{\hat{p}}}} \geq \frac{0.4-\mu_{\hat{p}}}{\sqrt{\sigma^2_{\hat{p}}}})$$

$$=P(\frac{\hat{p}-0.35}{\sqrt{0.0011375}} \geq \frac{0.4-0.35}{\sqrt{0.0011375}})$$

$$=P(Z \geq \frac{0.05}{0.0337})=P(Z \geq 1.48)$$

$$=0.5-P(0 \leq Z \leq 1.48)$$

$$=0.5-0.4306=0.0694$$

6 - 5 توزيع الفرق ما بين نسبتي عينتين

Distribution of the Difference Between Two Sample Proportions

هناك العديد من التطبيقات التي نرغب فيها معرفة الفرق ما بين نسبتي عينتين تم اختيارهم من مجتمعين إحصائيين ، وحساب الاحتمال المصاحب لهذا الفرق ، فمثلاً إذا كان هناك خط إنتاج بأحد المصانع وكان خط الإنتاج الأول ينتج سلع معينة بمعدل p_1 بينما خط الإنتاج الثاني ينتج سلع معينة بمعدل p_2 ، فمن الممكن أن ترغب إدارة المصنع في معرفة توزيع المعاينة للفرق ما بين نسبتي السلع المعيبة الموجودة في عينتين تم اختيارهما من خطي الإنتاج ، أو مثلاً من الممكن شركة ما تنتج نوعين من الروائح من الممكن شرائهما من قبل الرجال والنساء ، فإذا كانت p_1 تمثل نسبة الرجال الذين يفضلون شراء النوع الأول و p_2 تمثل نسبة الرجال الذين يفضلون شراء النوع الثاني ، فمن الممكن لإدارة الشركة اختيار عينة عشوائية من الرجال الذين يفضلون شراء النوع الأول وعينة من الرجال الذين يفضلون شراء النوع الثاني ومن هاتين العينتين يمكن تحديد في ما إذا كان هناك فرق ما بين نسبتي العينتين ، وحساب الاحتمال المصاحب لهذا الفرق ، وللقيام بهذا الإجراء يتطلب الأمر معرفة توزيع المعاينة (التوزيع الاحتمالي) للفرق ما بين نسبتي العينتين ، إن هذا التوزيع مضمون النظرية التالية والتي تعتبر امتداد للنظرية التي أشرنا إليها في حالة عينة واحدة .

نظرية (4) :- إذا كانت $X_1 , X_2 , X_3 , \dots , X_m$ تشكل عينة عشوائية من مجتمع إحصائي نسبة عناصرها التي لها الخاصية مدار البحث تساوى p_1 ونسبة عناصر العينة التي لها نفس الخاصية ($\hat{p}_1$) ، وكانت $Y_1 , Y_2 , Y_3 , \dots , Y_n$ تشكل عينة عشوائية من مجتمع إحصائي نسبة عناصره التي لها الخاصية مدار البحث تساوى p_2 ونسبة عناصر العينة التي لها نفس الخاصية ($\hat{p}_2$) وكانت العينتين مستقلتين وحجمها كبيراً فإن توزيع المعاينة للفرق ما بين نسبتي العينتين ($\hat{p}_1 - \hat{p}_2$) يتبع التوزيع الطبيعي تقريبياً بمتوسط يساوى $\mu_{\hat{p}_1 - \hat{p}_2} = p_1 - p_2$ وتباين $\sigma^2_{\hat{p}_1 - \hat{p}_2} = \frac{p_1(1-p_1)}{m} + \frac{p_2(1-p_2)}{n}$.

ملحوظة :

1 - إن شرط تطبيق النظرية السابقة هو أن تكون كلا من m و n كبيرة ، ويعتبر حجم العينة كبيراً إذا كان كلا من mp_1 و $m(1-p_1)$ و np_2 و $n(1-p_2)$ أكبر من 5 .

2- لحساب الاحتمالات المتعلقة بالفرق ما بين نسبتي العينتين تستخدم الصيغة الآتية :

$$Z = \frac{(\hat{p}_1 - \hat{p}_2) - \mu_{\hat{p}_1 - \hat{p}_2}}{\sqrt{\sigma^2_{\hat{p}_1 - \hat{p}_2}}}$$

حيث Z لها تقريباً توزيع طبيعي معياري .

مثال (10) : إذا علمت أنه يوجد نوعين من المبيدات الحشرية بالسوق ، وتدعى الشركة المصنعة للنوع الأول أنه يقضى على 90 ٪ من الحشرات عند استعماله وتدعى الشركة المصنعة للنوع الثاني أنه يقضى على 80 ٪ من الحشرات عند استعماله فإذا تم رش حجرتين لهما نفس الحجم بالنوع الأول والثانية بالنوع الثاني وتم اختيار عينتين من الحشرات حجم كل منها 200 ووضعت كل عينة في حجرة فما احتمال :

أ - أن يكون الفرق ما بين نسبتي العينتين أكبر من 20 ٪ ؟

ب - أن يكون الفرق ما بين نسبتي العينتين ما بين 15 ٪ و 18 ٪ ؟

الحل :

بفرض أن $p_1 = 0.90$ تمثل نسبة الحشرات التي يتم القضاء عليها باستخدام المبيد الأول .

$p_2 = 0.80$ تمثل نسبة الحشرات التي يتم القضاء عليها باستخدام المبيد الثاني .

و $\hat{p}_1$ تمثل نسبة الحشرات التي سيتم القضاء عليها بالعينة الأولى باستخدام المبيد الأول و $m = 200$.

و $\hat{p}_2$ تمثل نسبة الحشرات التي سيتم القضاء عليها بالعينة الثانية باستخدام المبيد الثاني و $n = 200$.

حيث أن حجم العينتين كبيراً وإن mp_1 و np_2 أكبر من 5 وعليه من النظرية السابقة فإن توزيع المعاينة للفرق ما بين نسبتي العينتين $(\hat{p}_1 - \hat{p}_2)$ يتبع التوزيع الطبيعي بمتوسط

$$\mu_{\hat{p}_1 - \hat{p}_2} = p_1 - p_2 = 0.90 - 0.80 = 0.10$$

وتباين

$$\sigma^2_{\hat{p}_1 - \hat{p}_2} = \frac{p_1(1-p_1)}{m} + \frac{p_2(1-p_2)}{n}$$
$$= \frac{(0.9)(0.10)}{200} + \frac{(0.8)(0.2)}{200} = \frac{0.25}{200} = 0.00125$$

أ -

$$P(\hat{p}_1-\hat{p}_2 \geq 0.20)=P\left(\frac{(\hat{p}_1-\hat{p}_2)-\mu_{\hat{p}_1-\hat{p}_2}}{\sqrt{\sigma^2_{\hat{p}_1-\hat{p}_2}}} \geq \frac{0.20-\mu_{\hat{p}_1-\hat{p}_2}}{\sqrt{\sigma^2_{\hat{p}_1-\hat{p}_2}}}\right)$$

$$=P\left(\frac{(\hat{p}_1-\hat{p}_2)-0.10}{\sqrt{0.00125}} \geq \frac{0.20-0.10}{\sqrt{0.00125}}\right)$$

$$=P\left(Z \geq \frac{0.10}{0.0354}\right)=P(Z \geq 2.82)$$

$$=0.5-P(0 \leq Z \leq 2.82)$$

$$=0.5-0.4976=0.0024$$

أي أن هناك احتمال قدرة 0.24 ٪ بأن يكون هناك فرق ما بين نسبتي عينتي الحشرات التي سيتم القضاء عليها باستخدام نوعي المبيد .

ب -

$$P(0.15 \leq \hat{p}_1-\hat{p}_2 \leq 0.18)=P\left(\frac{0.15-\mu_{\hat{p}_1-\hat{p}_2}}{\sqrt{\sigma^2_{\hat{p}_1-\hat{p}_2}}} \leq \frac{(\hat{p}_1-\hat{p}_2)-\mu_{\hat{p}_1-\hat{p}_2}}{\sqrt{\sigma^2_{\hat{p}_1-\hat{p}_2}}} \leq \frac{0.18-\mu_{\hat{p}_1-\hat{p}_2}}{\sqrt{\sigma^2_{\hat{p}_1-\hat{p}_2}}}\right)$$

$$=P\left(\frac{0.15-0.10}{\sqrt{0.00125}} \leq \frac{(\hat{p}_1-\hat{p}_2)-0.10}{\sqrt{0.00125}} \leq \frac{0.18-0.10}{\sqrt{0.00125}}\right)$$

$$=P\left(\frac{0.05}{0.0354} \leq Z \leq \frac{0.08}{0.0354}\right)$$

$$=P(1.41 \leq Z \leq 2.26)$$

$$=P(0 \leq Z \leq 2.26)-P(0 \leq Z \leq 1.41)$$

$$=0.4881-0.4207=0.0674$$

اي أن احتمال أن يكون الفرق ما بين نسبتي عينتي الحشرات التي سيتم القضاء عليها باستخدام نوعي المبيد ما بين 0.15 و 0.18 يساوي 6.74 ٪ .

مثال (**11**) : إذا علمت أن الشركة العامة للإلكترونات تستورد قطع الغيار اللازمة لها يومياً من مصدرين مختلفين وإن القطع الواردة إليها من المصدر الأول ترفض بمعدل 8 ٪ بسبب عيوب بها ، بينما القطع الواردة إليها من المصدر الثاني ترفض بمعدل 5 ٪ لنفس السبب ، فإذا كانت خطوط الإنتاج تستخدم 150 قطعة من المصدر الأول و 300 قطعة من المصدر الثاني

يومياً، فما هي نسبة الأيام التي سوف يكون فيها الفرق ما بين نسبتي الرفض من المصدرين أقل من أو يساوى 7 ٪ ؟

الحل :

بفرض أن $\hat{p}_1$ تمثل نسبة القطع المرفوضة يومياً بالعينة العشوائية المختارة من المصدر الأول ، و $\hat{p}_2$ تمثل نسبة القطع المرفوضة يومياً بالعينة العشوائية المختارة من المصدر الثاني .

وحيث أن الاستخدام اليومي للقطع المستوردة تتضمن عينات عشوائية كبيرة من المصدرين وعليه فإن الفرق ما بين نسبتي العينتين $(\hat{p}_1 - \hat{p}_2)$ لكل يوماً يمكن تقريبه باستخدام التوزيع الطبيعي ، بمتوسط يساوى $\mu_{\hat{p}_1 - \hat{p}_2} = p_1 - p_2 = 0.08 - 0.05 = 0.03$ وتباين يساوى

$$\sigma^2_{\hat{p}_1 - \hat{p}_2} = \frac{p_1(1-p_1)}{m} + \frac{p_2(1-p_2)}{n}$$

$$= \frac{(0.08)(0.92)}{150} + \frac{(0.05)(0.95)}{300} = 0.00065$$

إذن

$$P((\hat{p}_1 - \hat{p}_2) \le 0.01) = P(\frac{(\hat{p}_1 - \hat{p}_2) - \mu_{\hat{p}_1 - \hat{p}_2}}{\sqrt{\sigma^2_{\hat{p}_1 - \hat{p}_2}}} \le \frac{0.07 - \mu_{\hat{p}_1 - \hat{p}_2}}{\sqrt{\sigma^2_{\hat{p}_1 - \hat{p}_2}}})$$

$$= P(\frac{(\hat{p}_1 - \hat{p}_2) - 0.03}{\sqrt{0.00065}} \le \frac{0.07 - 0.03}{\sqrt{0.00065}})$$

$$= P(Z \le \frac{0.04}{0.0255})$$

$$= P(Z \le 1.57) = 0.9418$$

أي أنه حوالي 94 ٪ من الأيام سيكون فيها الفرق ما بين نسبتي الرفض من المصدرين أقل من أو يساوى 7 ٪ .

مثال (12) : إذا علمت أن 40 ٪ من مجتمع الأطفال المعاقين (A) قادرين على الحركة ، وأن 40 ٪ من مجتمع الأطفال المعاقين (B) قادرين على الحركة ، وتم اختيار عينة عشوائية تتكون من 120 طفل من المجتمع (A) واختيرت عينة عشوائية تتكون من 100 طفل من المجتمع (B) . فما احتمال أن يكون الفرق ما بين نسبتي العينتين 16 ٪ أو أكثر ؟

الحل :

بفرض أن $p_1 = 0.40$ تمثل نسبة الأطفال المعاقين و القادرين على الحركة بالمجتمع (A) .

$p_2 = 0.40$ تمثل نسبة الأطفال المعاقين والقادرين على الحركة بالمجتمع (B) .

وإن $\hat{p}_1$ تمثل نسبة الأطفال المعاقين والقادرين على الحركة بعينة المجتمع (A) .

و $\hat{p}_2$ تمثل نسبة الأطفال المعاقين والقادرين على الحركة بعينة المجتمع (B) .

حيث أن حجم العينتين كبيراً وعليه يمكن الافتراض بأن توزيع المعاينة للفرق ما بين نسبتي العينتين $(\hat{p}_1 - \hat{p}_2)$ يتوزع وفق التوزيع الطبيعي تقريبياً بمتوسط يساوي :

$$\mu_{\hat{p}_1-\hat{p}_2} = p_1 - p_2 = 0.40 - 0.40 = 0$$

وتباين يساوي :

$$\sigma^2_{\hat{p}_1-\hat{p}_2} = \frac{p_1(1-p_1)}{m} + \frac{p_2(1-p_2)}{n} = \frac{(0.40)(1-0.40)}{120} + \frac{(0.40)(1-0.40)}{100}$$
$$= 0.002 + 0.0024 = 0.0044$$

وعليه فإن

$$P((\hat{p}_1 - \hat{p}_2) \geq 0.16) = P\left(\frac{(\hat{p}_1 - \hat{p}_2) - \mu_{\hat{p}_1-\hat{p}_2}}{\sqrt{\sigma^2_{\hat{p}_1-\hat{p}_2}}} \geq \frac{0.16 - \mu_{\hat{p}_1-\hat{p}_2}}{\sqrt{\sigma^2_{\hat{p}_1-\hat{p}_2}}}\right)$$
$$= P\left(\frac{(\bar{p}_1 - \hat{p}_2) - 0}{\sqrt{0.0044}} \geq \frac{0.16 - 0}{\sqrt{0.0044}}\right)$$
$$= P\left(Z \geq \frac{0.16}{0.0663}\right) = P(Z \geq 2.41)$$
$$= 0.5 - 0.4920 = 0.008$$

1- إذا علمت أن الإنفـاق الأسـبوعي لأفـراد مجتمـع مـا يتـوزع وفـق التوزيـع الطبيعـي بمتوسـط يساوى 49.79 دينار وبانحراف معياري يساوى 11.72 دينار فإذا تم اختيار عينة عشـوائية مـن 10 أفراد فما احتمال :

أ - أن يكون متوسط الإنفاق الأسبوعي ما بين 47.79 و 51.79 .

ب - أن يكون متوسط الإنفاق الأسبوعي ما بين 44.79 و 54.79 .

جـ - أن يكون متوسط الإنفاق الأسبوعي ما بين 49.64 و 49.94 .

2- يدعى أحد التجار المتخصصين في بيع النضائد السائلة أن مبيعاته خلال شهر معين من النضائد التي أعمارها الافتراضية سنة وسنتين وثلاث سنوات كانت متساوية فإذا تم اختيار عينة عشوائية حجمها 2 من هذه النضائد فأوجد :

أ - جميع العينات الممكنة الاختبار في حالة مع الإعادة وبدون إعادة .

ب - أوجد التوزيع الاحتمالي لمتوسط العينات الممكنة السحب بالفقرة " أ " ثم مثله بيانياً .

جـ - جميع العينات الممكنة الاختيار إذا كانت $n = 3$ ومع الإعادة ثم أوجد التوزيع الاحتمالي لمتوسط هذه العينات .

3- إذا علمت أن أوزان علب السردين المنتجة من أحد المصانع تتوزع وفق التوزيع الطبيعي بمتوسط يساوى 32.1 جرام وبانحراف معياري يساوى 0.2 جرام ، فما احتمال أن يكون متوسط العينة أقل من 32 إذا اشتريت

أ - علبة واحدة　　ب - 4 علب　　جـ - 16 علبة .

4 - تستخدم آلة معينة لتعبئة قوارير المشروبات الغازية التى سعتها 500 مل ، فإذا وضعت هذه الآلة تحت المراقبة لفترة زمنية معينة فوجد أن التباين في الكمية المعبأة باستخدام هذه الآلة يساوى 1.0 مل ، واختيرت عينة عشوائية تتكون من 25 قاروره ، وعلى افتراض أن مجتمع المعاينة مجتمع طبيعي تقريباً أوجد :

أ- $P\left[\left|\bar{X}-\mu\right| \leq 0.3\right]$

5- إذا علمت أن ضغط الدم للأشخاص الأصحاء يتوزع وفق التوزيع الطبيعي بمتوسط يساوي 120 وبانحراف معياري يساوى 10 فإذا تم اختيار عينة عشوائية تتكون من 12 شخص فما احتمال

أ - أن يكون متوسط ضغط الدم لأفراد هذه العينة أقل من 124 .

ب - أن يكون متوسط ضغط الدم لأفراد هذه العينة أكبر من 115 .

جـ - أن يكون متوسط ضغط الدم لأفراد هذه العينة ما بين 122 و 125 .

6 - إذا علمت أن أوزان أكياس الدقيق المنتجة من قبل أحد المصانع تتوزع وفق التوزيع الطبيعي بمتوسط يساوى 100 كجم وتباين يساوى 64 كجم ، فإذا تم اختيار عينة عشوائية تتكون من 16 كيس فما احتمال

أ - أن يكون متوسط أوزان الأكياس بهذه العينة أكبر من 106 كجم .

ب - أن يكون متوسط أوزان الأكياس ما بين 95 كجم و 99 كجم .

7- إذا كانت جميع العينات الممكنة ذات الحجم 16 تم اختيارها من مجتمع طبيعي بمتوسط يساوى 50 وبانحراف معياري يساوى 5 ، فما احتمال أن يقع متوسط العينة $(\bar{X})$ ما بين $\mu - 1.96\sigma_{\bar{x}}$ و $\mu + 1.96\sigma_{\bar{x}}$.

8 - مصنع لإنتاج النضائد السائلة يدعى أن أعمار البطاريات التي ينتجها يتوزع وفق التوزيع الطبيعي بمتوسط يساوى 54 شهراً و بانحراف معياري يساوى 7 أشهر ، فإذا تم اختيار عينة عشوائية تتكون من 50 نضيدة من إنتاج هذا المصنع ووضعت تحت الاختبار وذلك للتأكد من صحة ادعاء المصنع ، فما احتمال

أ - أن يكون متوسط أعمار البطاريات التي بالعينة أقل من 52 شهراً .

ب - أن يكون متوسط أعمار البطاريات التي بالعينة ما بين 56 و 57 شهراً .

جـ - أن يكون متوسط أعمار البطاريات التي بالعينة أقل من 57 شهراً .

د - أن يكون متوسط أعمار البطاريات التي بالعينة أكبر من 53 شهراً .

9 - إذا علمت أن متوسط أجرة ساعة العمل بأحد المواقع الصناعية يساوى 8.75 دينار وبانحراف معياري يساوى 0.37 دينار ، وإن متوسط أجرة ساعة العمل بموقع صناعي أخر يساوى 7.92 دينار وبانحراف معياري يساوى 0.86 دينار ، فإذا تم اختيار عينة عشوائية حجمها 100 عامل من كل موقع فما احتمال

أ - أن يكون الفرق ما بين متوسطي أجرة ساعة العمل بالعينتين أكثر من دينار واحد .

ب - أن يكون الفرق ما بين متوسطي أجرة ساعة العمل بالعينتين ما بين 0.75 دينار ودينار واحد .

جـ - أن يكون الفرق ما بين متوسطي أجرة ساعة العمل بالعينتين أقل من 1.10 دينار .

10- أجريت دراسة حول الإنفاق العائلي السنوي (بالدينار) على الترفيه ، وذلك من خلال دراسة مجتمعين فكانت النتائج كما يلي :

المجتمع I : $m = 40 \qquad \mu_1 = 332 \qquad \sigma_1^2 = 2860$

المجتمع II : $n = 35 \qquad \mu_2 = 300 \qquad \sigma_2^2 = 3250$

فما احتمال :

أ - أن يكون الفرق ما بين متوسطي الإنفاق بالعينتين أكبر من 50 دينار .

ب - أن يكون الفرق ما بين متوسطي الإنفاق بالعينتين أقل من 18 دينار .

جـ - أن يكون الفرق ما بين متوسطي الإنفاق بالعينتين ما بين 50 و 55 دينار .

11- إذا علمت أن قراءات مستويات مصل الكولسترول لمجموعتين من الأعمار في بلد ما كما يلي :

المجموعة	العمر	المتوسط	الانحراف المعياري
الأولى	18 - 22	179	42
الثانية	23 - 30	198	48

فإذا تم اختيار عينتين عشوائيتين مستقلتين حجمها كل منهما 50 فما احتمال

أ - أن يكون الفرق ما بين متوسطي العينتين أقل من 23 .

. أكبر من 25

ب – أن يكون الفرق ما بين متوسطي العينتين ما بين 25 و 27 .

جـ – أن يكون الفرق ما بين متوسطي العينتين

12- إذا علمت انه تم اختيار عينة عشوائية حجمها 100 من مجتمع ما و إن $P = 0.6$ أوجد :

أ – $P(\hat{p} \leq 0.58)$.

ب – $P(\hat{p} \geq 0.65)$.

جـ – $P(0.56 \leq \hat{p} \leq 0.63)$.

13- إذا علمت أن 32.2 ٪ من النساء اللواتي أعمارهن 35 سنة أو أكثر أجريت لهن فحوصا على سرطان الثدي خلال السنة الماضية ، فإذا تم اختيار عينة عشوائية تتضمن 150 امرأة من هذا المجتمع فما احتمال :

أ – أن يكون نسبة النساء اللواتي أجرى لهن فحوصا بهذه العينة ما بين 0.26 و 0.35 ؟

ب – أن يكون نسبة النساء اللواتي أجرى لهن فحوصا بهذه العينة أكثر من 0.28 ؟

جـ – أن يكون نسبة النساء اللواتي أجرى لهن فحوصا بهذه العينة أقل من 0.36 ؟

14- إذا علمت أن نسبة الأفراد الذين يتعاطون المخدرات بمجتمع ما يساوى 0.45 ، وإن نسبة الأفراد الذين يتعاطون المخدرات في مجتمع أخر يساوى 0.28 ، فإذا تم اختيار عينتين عشوائيتين مستقلتين حجم كل منهما 90 فما احتمال :

أ – أن يكون نسبة الفرق ما بين نسبتي العينتين أكبر من 0.25 .

ب – أن يكون نسبة الفرق ما بين نسبتي العينتين أقل من 0.08 .

جـ – أن يكون نسبة الفرق ما بين نسبتي العينتين 0.07 و 0.26 .

15- إذا علمت أن 14 ٪ من الذكور و 24 ٪ من الإناث الذين أعمارهم ما بين 25 و 65 سنة يعانون من زيادة في الوزن ، فإذا تم اختيار عينة عشوائية من الذكور حجمها 100 وعينة عشوائية من الإناث حجمها 110 وكانت العينتين مستقلتين فأوجد :

أ – احتمال أن يكون الفرق ما بين نسبتي العينتين ما بين 0.02 و 0.16 .

ب – احتمال أن يكون الفرق ما بين نسبتي العينتين أقل من 0.18 .

جـ – احتمال أن يكون الفرق ما بين نسبتي العينتين أكبر من 0.04 .

د - احتمال أن يكون الفرق ما بين نسبتي العينتين ما بين 0.02 و 0.04 .

16- إذا علمت أن 10 ٪ من القضايا المرفوعة ضد شركات التأمين يكون فيها الحكم لصالح المدعى ، فإذا تم اختيار عينة عشوائية من 50 قضية من القضايا المرفوعة ضد شركات التأمين فما احتمال :

أ - أن تكون نسبة القضايا التي يصدر فيها الحكم لصالح المدعى لا تزيد عن 15 ٪ .

ب -أن تكون نسبة القضايا التي يصدر فيها الحكم لصالح المدعى ما بين 16 ٪ و 20 ٪ .

17- إذا علمت أن أوزان حقائب المسافرين على متن الخطوط الجوية العربية الليبية تتوزع وفق التوزيع الطبيعي بمتوسط يساوى 20 كجم وبانحراف معياري يساوى 4 كجم ، فإذا تم اختيار عينة عشوائية من 25 حقيبة فما احتمال :

أ - أن يكون متوسط أوزان الحقائب بهذه العينة لا يقل عن 18 كجم .

ب - أن يكون متوسط أوزان الحقائب بهذه العينة ما بين 18 و 22 كجم .

18- إذا علمت أنه تم اختيار عينة عشوائية حجمها 25 من مجتمع طبيعي بمتوسط يساوى 8 وتباين يساوى 16 أوجد :

أ- $P(\overline{X} \le 10)$.

ب - $P(|\overline{X}| \ge 9)$.

جـ - $P(|\overline{X} - \mu| \le 1)$.

19- إذا علمت أن العمر الزمني الذي تعمره المصابيح الكهربائية التي من النوع 75 وات يتوزع وفق التوزيع الطبيعي تقريبياً بمتوسط يساوى 1014 ساعة وبانحراف معياري يساوى 25 ساعة ، فإذا تم اختيار عينة عشوائية تتكون من 20 مصباحاً فما احتمال :

أ- أن يكون متوسط العمر الزمن لمصابيح العينة أكبر من 1015 .

ب - أن يكون متوسط العمر الزمن لمصابيح العينة أقل من 1012 .

جـ - أن يكون متوسط العمر الزمن لمصابيح العينة ما بين 1010 ساعة و 1016 ساعة .

20- يفرض أنه تم اختيار عينة عشوائية حجمها 50 من مجتمع بمتوسط يساوي 50 والانحراف معياري يساوي 2 أوجد :

أ- $P(\overline{X} \geq 51)$.

ب- $P(|\overline{X} - 50| < 0.5)$.

جـ- أوجد قيمة الثابت a بحيث $P(\overline{X} < a) = 0.05$.

د- أوجد قيمة الثابت a بحيث $P(|\overline{X} - 50| < a) = 0.90$.

21- إذا علمت أن 10 ٪ من المقالات المكتوبة في الصحف اليومية بها أخطاء لغوية ، فإذا تم اختيار عينة عشوائية من 100 مقال فأوجد :

أ- احتمال أن تكون نسبة الأخطاء بها على الأكثر 12 ٪ .

ب- احتمال أن يكون نسبة الأخطاء بها على الأقل 15 ٪ .

جـ- احتمال أن يكون نسبة الأخطاء بها ما بين 8 ٪ و 13 ٪ .

22- إذا تم اختيار عينة عشوائية حجمها 5 من التوزيع المنتظم المعرف على الفترة (1 , 0) ، فأوجد :

أ- $P(0.25 < \overline{X} < 0.75)$.

ب- $P(\overline{X} \geq 0.65)$.

جـ- $P(\overline{X} \leq 0.35)$.

23- إذا كانت $X_1, X_2, \ldots, X_n$ عينة عشوائية من مجتمع طبيعي بمتوسط يساوي μ وتباين يساوي 100 حيث μ غير معلومة ، أوجد :

أ- $P\left[|\overline{X} - \mu| \leq 0.1\right]$ عندما $n = 9$.

ب- $P\left[|\overline{X} - \mu| \leq 0.1\right]$ عندما $n = 36$.

24- إذا علمت أن 50 ٪ من الأشخاص المتبرعين بالدم بالمدينة (A) فصيلة دمهم (O^+) و33 ٪ من الأشخاص المتبرعين بالدم بالمدينة (B) فصيلة دمهم (O^+) ، فإذا تم اختيار عينتين عشوائيتين مستقلتين من الفصيلتين حجم كل منهما 100 ، فأوجد :

أ- احتمال أن يكون الفرق ما بين نسبتي العينتين أكبر من 30 ٪ .

ب - احتمال أن يكون الفرق ما بين نسبتي العينتين أقل من 20 ٪ .

جـ - احتمال أن يكون الفرق ما بين نسبتي العينتين ما بين 10 ٪ و 25 ٪ .

25- إذا افترضنا أن نسب الفاقد في طلاب المرحلة الثانويـة بمدينتيـن هـي 25 ٪ و 15 ٪ علـى التوالي وإذا تم اختيار عينتين عشوائيتين مستقلتين حجم كل منها 50 ، فأوجد :

أ- احتمال أن يكون الفرق ما بين نسبتي العينتين أقل من 20 ٪ .

ب - احتمال أن يكون الفرق ما بين نسبتي العينتين ما بين 18 ٪ و 25 ٪ .

الفصل التـــاسع

الانحدار والارتبــــــاط

Regression and Correlation

9-1 مقدمــــــــة Introduction

إن العديد من الأبحاث الإحصائية يكون الهدف منها هو البحث في إمكانية إيجاد علاقة ما بين متغيرين أو أكثر فمثلاً دراسة :

أ - العلاقة ما بين كمية المبيعات من سلعة ما وسعرها .

ب - العلاقة ما بين الوزن والطول .

جـ - العلاقة ما بين الكمية المطلوبة من سلعة معينة وسعرها وأسعار السلع المنافسة لها ودخل الفرد والحالة الاجتماعية .

د - العلاقة ما بين ضغط الدم والعمر .

هـ - العلاقة ما بين كمية الكولسترول في الدم والوزن والعمر والطول وعادة التدخين .

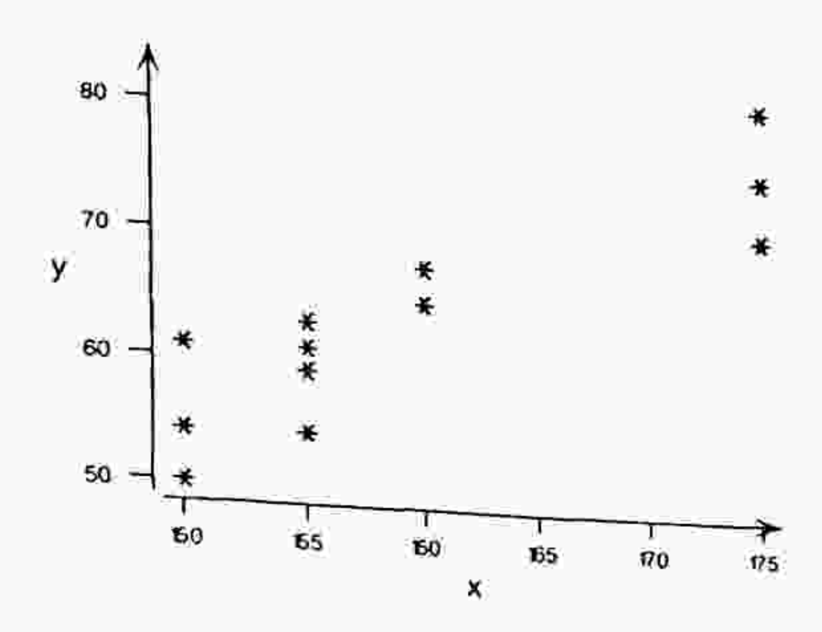
و - العلاقة ما بين سرعة الرياح وارتفاعها عن الأرض ودرجة الحرارة والضغط الجوى .

إن تحليل الانحدار هو أسلوب يستخدم في تحديد نوعية العلاقة ما بين متغيرين أو أكثر ثم استخدام هذه العلاقة في التنبؤ أو تقدير قيمة متغير ما بمعرفة قيمة المتغير أو المتغيرات الأخرى ، إن المعادلة الرياضية التى يمكن من خلالها التنبؤ بقيم متغير ما من خلال معرفة قيم متغير آخر أو أكثر تسمى بمعادلة الانحدار (regression equation) وقد تكون هذه المعادلة خطية أو غير خطية ، وإذا تضمنت هذه المعادلة متغيرين فقط فإنها تسمى معادلة انحدار بسيط أما إذا كان هناك أكثر من متغيرين فإنها تسمى معادلة انحدار متعدد .

ولتطبيق تحليل الانحدار عن أى ظاهرة يجب معرفة أو افتراض طبيعة العلاقة ما بين المتغيرات قيد الدراسة ، ويتم ذلك من خلال التعبير عن تلك العلاقة بمعادلة رياضية ، فإذا افترضنا أنه تم اختيار عينة عشوائية من 12 شخص وتم قياس طول ووزن كل منهم وكانت النتائج كما يلي :

الشخص	1	2	3	4	5	6	7	8	9	10	11	12
الطول	150	150	150	155	155	155	155	160	160	175	175	175
الوزن	50	61	54	54	63	59	61	68	65	77	83	72

إن أفضل طريقة لبداية تحليل مثل هذا النوع من البيانات باستخدام أسلوب الانحدار هو رسم الشكل الإنتشارى لهذه البيانات وذلك بهدف التعرف على طبيعة العلاقة ما بين هذين المتغيرين كما يلي :



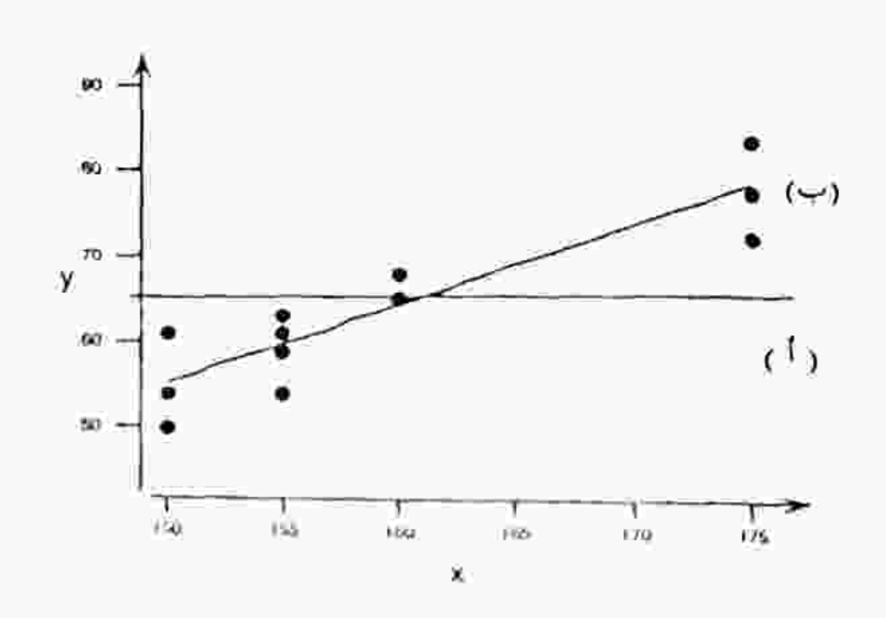
شكل (1) : الشكل الأنتشارى لبيانات الجدول السابق .

وبرسم خط خلال هذه النقاط يتبين لنا وجود علاقة ما بين الوزن والطول .

فإذا افترضنا أن قيم الأطوال قد تم اختيارها قبل جمع البيانات ثم صنف الأشخاص حسب الأطوال ثم اختيرت عينة عشوائية من كل طول فإنه يمكن اعتبار هذه البيانات قد تم الحصول عليها من تجربة تتضمن أربعة مجتمعات كما يتضح في الجدول الآتي :

الأطوال	الأوزان	المجموع	المتوسط
150	54 61 50	165	μ_{150} = 55
155	61 59 63 54	237	μ_{155} = 59.25
160	65 68	133	μ_{160} = 66.5
175	72 83 77	232	μ_{175} = 77.333

ومن هذا الجدول يتضح أنه هناك أربعة متوسطات μ_{150} ، μ_{155} ، μ_{160} ، μ_{175} إن هذه المتوسطات تعتمد على قيمة الطول ، وهذا يعني أن متوسط الأوزان من الممكن أن يختلف من طول إلى آخر .وعليه يمكن القول بأن توزيع الأوزان سوف يعتمد على قيم الأطوال التي تم اختيارها ، وعليه يطلق على الأوزان في هذه الحالة تسمية المتغير التابع وسوف نرمز لهذا المتغير بالرمز Y ويطلق على الأطوال تسمية المتغير المستقل وسوف نرمز لهذا المتغير بالرمز X .وبالنظر إلى الشكل البياني (2) نلاحظ أن انحرافات القيم عن الخط " ب " أصغر من انحرافات القيم عن الخط " أ " أي أن مجموع مربع انحرافات القيم عن الخط " ب " تعطي تقديراً أصغر لتباين الأوزان ، وبعبارة أخرى إن التباين حول الخط " ب " أصغر بكثير من التباين حول الخط (أ) أي التباين حول متوسط الأوزان ، وعليه يمكن القول بإضافة معلومات حول الطول لدراسة الوزن يمكن الحصول على معلومات أكثر حول الوزن ، أي أن بعض الاختلاف غير المفسر للأوزان يمكن تفسيره من خلال معرفة الأطوال .



شكل (2) : انحرافات القيم عن المتوسط والخط المستقيم .

تعريف (1) : إن انحدار Y على X هو متوسط توزيع Y عند قيمة معينة للمتغير X ويرمز لذلك بالرمز $\mu_{Y/X=x}$ أو ببساطة $\mu_{Y/x}$.

تعريف (2) : دالة الانحدار (Regression function) هي الدالة الافتراضية ما بين $\mu_{Y/x}$ و X أي أن $\mu_{Y/x}$ يفترض بأن تكون دالة معروفة في X ، وبالمفهوم الرياضي $\mu_{Y/x} = f(x)$.

الحظ أنه من الممكن أن يكون هناك أكثر من متغير مستقل واحد وسيكون متوسط Y في هذه الحالة معتمداً على جميع قيم المتغيرات المستقلة وبالتالي تكون دالة الانحدار على الصورة الآتية:

$$\mu_{Y/x_1,x_2,x_3,\ldots,x_k} = f(x_1, x_2, x_3, \ldots, x_k)$$

9 – 2 الانحدار الخطي البسيط Simple linear regression

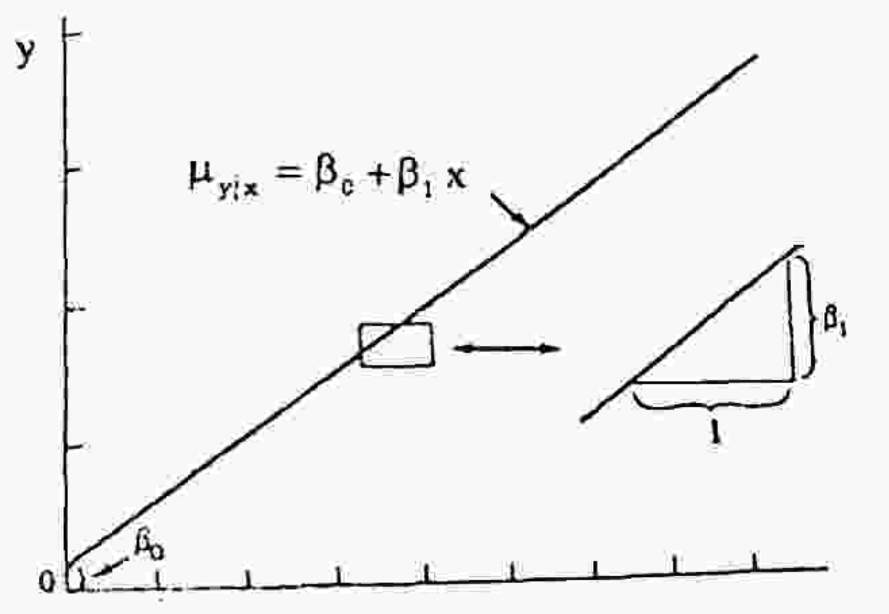
إن ابسط نماذج الانحدار هو نموذج الانحدار الخطي البسيط ويقوم هذا النموذج على الافتراضات الآتية :

1 – إن قيم المتغير المستقل X بالعينة العشوائية $\{(X_i , Y_i) , i = 1,2,3,\ldots,n\}$ تكون ثابتة وهذا يعني أن قيم X قد تم اختيارها مسبقاً من قبل الباحث وبالتالي عند جمع البيانات عن المتغير Y سوف لن تتغير تلك القيم .

2 – إن متوسطات توزيع Y $(\mu_{Y/x})$ عند كل قيمة من قيم X تقع على نفس الخط المستقيم هذه الفرضية يمكن كتابتها كما يلي :

$$\mu_{Y/x} = \beta_0 + \beta_1 x \qquad (1)$$

حيث $\mu_{Y/x}$ تمثل متوسط توزيع قيم Y لقيمة معينة من X ، و β_0 و β_1 يطلق عليهما تسمية معلمتي الانحدار وإن β_0 تمثل قيمة $\mu_{Y/x}$ عندما $X = 0$ و β_1 تمثل ميل الخط الذي تقع عليه جميع المتوسطات .

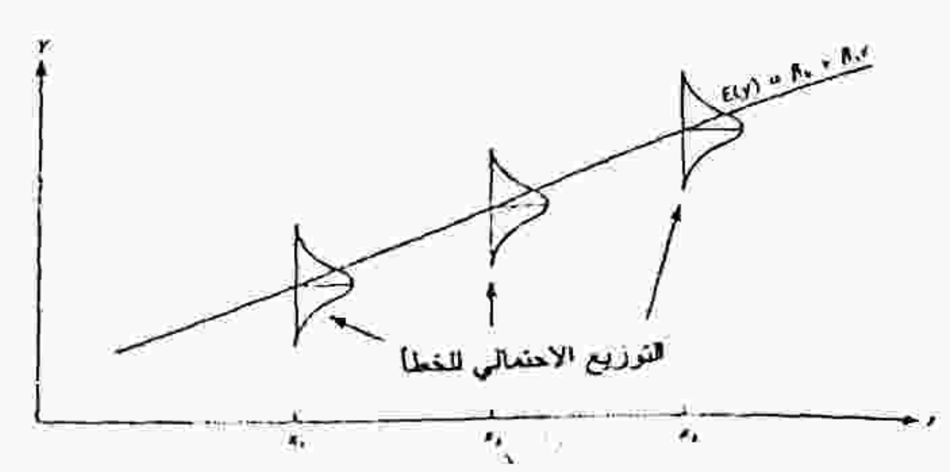


شكل (3) : مفهوم دالة الانحدار

3 – إن العلاقة ما بين قيم Y المشاهدة و $\mu_{Y/x}$ و X تكون نموذج الانحدار الذي يكون على الصورة الآتية :

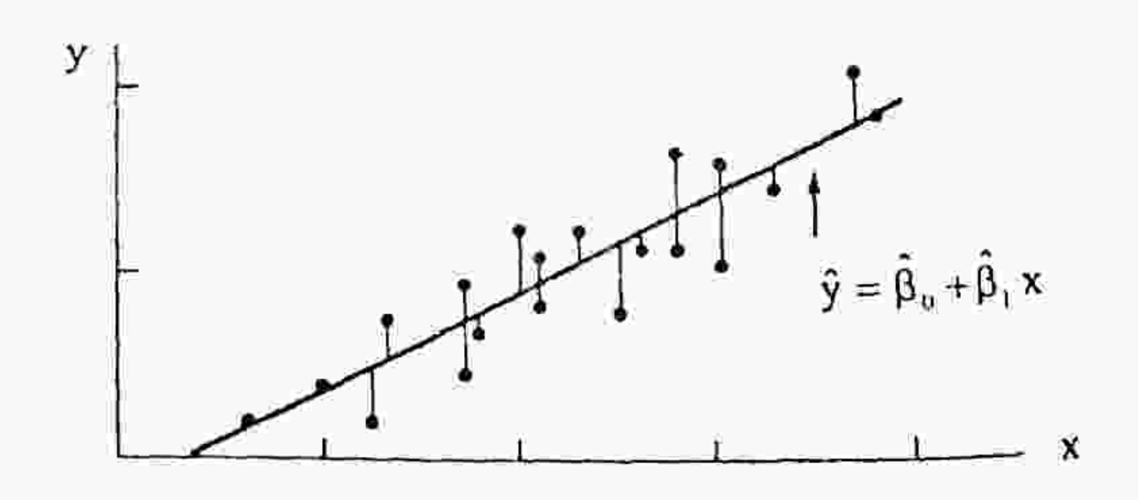
$$y_i = \mu_{Y/x} + \in_i$$
$$= \beta_0 + \beta_1 x_i + \in_i \quad , i = 1,2,3,\ldots,n \qquad (2)$$

حيث $\in_i$ يمثل الخطأ وله توزيع طبيعي بمتوسط يساوى صفر وتباين يساوى σ^2 وهو ثابت لكل قيمة من قيم X ، وعليه فإن تباين Y يساوى $\sigma^2_{Y/x}$ حيث $\sigma^2_{Y/x} = \sigma^2$ وله قيمة واحدة لأية قيمة من قيم X أيضاً.



شكل (4) :التوزيع الاحتمالي للخطأ $\in$: $\sigma_{Y/x_1} = \sigma_{Y/x_2} = \sigma_{Y/x_3} = \cdots = \sigma_{Y/x_n}$.

حيث أن $\mu_{Y/x}$ تعتمد على قيمتي β_0 و β_1 وكل منهما معلمة مجهولة وإن $\sigma^2_{Y/x}$ غير معلومة أيضاً وبالتالي يتم تقدير هذه المعلمات من واقع البيانات التى تم جمعها عن الظاهرة مدار البحث . إن طريقة المربعات الصغرى (Least square method) هي أكثر الطرق استخداماً في تقدير هذه المعلمات ، ووفقاً لهذه الطريقة يتم اختيار قيمتي β_0 و β_1 وليكونا $\hat{\beta}_0$ و $\hat{\beta}_1$ بحيث يكون مجموع مربع انحرافات المفردات عن خط الانحدار أصغر ما يمكن .



شكل (5) : انحرافات y_i عن خط الانحدار .

إن انحرافات y_i عن $\mu_{y/x}$ تكون كالآتي :

$$\epsilon_i = y_i - \mu_{Y/x_i} = y_i - \beta_0 - \beta_1 x_i$$

(3)

وبالتالي يتم اختيار $\hat{\beta}_0$ و $\hat{\beta}_1$ بحيث يكون

$$Q = \sum_{i=1}^{n} \epsilon_i^2 = \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)^2$$

اصغر ما يمكن ، ويمكن الحصول على ذلك من خلال تفاضل Q بالنسبة الى كـلا من β_0 و β_1 وتسوية الناتج بالصفر ثم حل المعادلتين الناتجتين ،أى أن :

$$\frac{\partial(Q)}{\partial \beta_0} = -2\sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i) = 0$$

$$\frac{\partial(Q)}{\partial \beta_1} = -2\sum_{i=1}^{n} x_i (y_i - \beta_0 - \beta_1 x_i) = 0$$

وعليه فإن :

$$n\hat{\beta}_0 + \hat{\beta}_1 \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i$$

$$\hat{\beta}_0 \sum_{i=1}^{n} x_i + \hat{\beta}_1 \sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i y_i \qquad (4)$$

وهاتين المعادلتين يطلق عليهما تسمية المعادلتين الطبيعيتين .

ومنهما نجد أن

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} (x_i - \overline{x})(y_i - \overline{y})}{\sum_{i=1}^{n} (x_i - \overline{x})^2} = \frac{SS_{xy}}{SS_{xx}} \qquad (5)$$

ألحظ أن SS_{xy} يمكن إعادة كتابتها كما يلي :

$$SS_{xy} = \sum_{i=1}^{n} x_i y_i - n\overline{x}\overline{y} = \sum_{i=1}^{n} x_i y_i - \frac{\left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right)}{n} \qquad (6)$$

و

$$SS_{xx} = \sum_{i=1}^{n} x^2 - \frac{(\sum_{i=1}^{n} x_i)^2}{n} = \sum_{i=1}^{n} x_i^2 - n(\overline{x})^2 \qquad (7)$$

وإن

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} \quad (8)$$

وعليه فإن تقدير $\mu_{Y/x}$ لقيمة معينة من قيم x يكون كالآتي :

$$\begin{aligned} \hat{y} &= \hat{\beta}_0 + \hat{\beta}_1 x \\ &= \hat{\beta}_0 + \hat{\beta}_1 \bar{x} + \hat{\beta}_1 x - \hat{\beta}_1 \bar{x} \\ &= \bar{y} + \hat{\beta}_1 (x - \bar{x}) \end{aligned} \quad (9)$$

ويمكن إيجاد هذه القيم التقديرية من بيانات المثال السابق وذلك كما يلي :

x_i	y_i	x_i^2	y_i^2	$x_i y_i$
150	50	22500	2500	7500
150	61	22500	3721	9150
150	54	22500	2916	8100
155	54	24025	2916	8370
155	63	24025	3969	8765
155	59	24025	3481	9145
155	61	24025	3721	9455
160	68	25600	4624	10880
160	65	25600	4225	10400
175	77	30625	5929	13475
175	83	30625	6889	14525
175	72	30625	5184	12600
$\sum_{i=1}^{12} x_i = 1915$	$\sum_{i=1}^{12} y_i = 767$	$\sum_{i=1}^{12} x_i^2 = 306675$	$\sum_{i=1}^{12} y_i^2 = 50075$	$\sum_{i=1}^{12} x_i y_i = 123365$

وعليه فإن $\bar{x} = 159.58$ و $\bar{y} = 63.92$

إذن

$$SS_{xy} = \sum_{i=1}^{12} x_i y_i - \frac{\left(\sum_{i=1}^{12} x_i\right)\left(\sum_{i=1}^{12} y_i\right)}{n} = 123365 - \frac{(1915)(767)}{12} = 964.583$$

$$SS_{xx} = \sum_{i=1}^{12} x_i^2 - \frac{\left(\sum_{i=1}^{12} x_i\right)^2}{n} = 306675 - \frac{(1915)^2}{12} = 1072.917$$

وإن

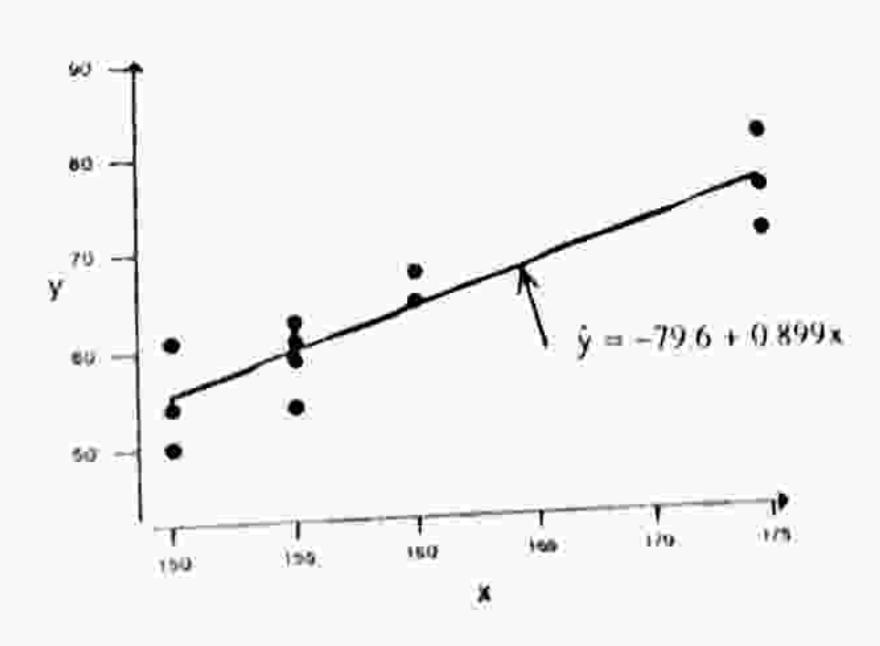
$$\hat{\beta}_1 = \frac{SS_{xy}}{SS_{xx}} = \frac{964.583}{1072.917} = 0.899$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} = 63.92 - (0.899)(159.58) \approx -79.6$$

وبالتعويض عن $\hat{\beta}_0$ و $\hat{\beta}_1$ في معادلة الانحدار التقديرية نجد أن :

$$\begin{aligned} \hat{y} &= \hat{\beta}_0 + \hat{\beta}_1 x \\ &= -79.6 + 0.899\,x \end{aligned}$$

ويمكن تمثيل معادلة الانحدار التقديرية $\hat{y}$ بيانياً كما يلي :



شكل (6) : الشكل الانتشارى للبيانات وخط معادلة الأنحدار التقديرية .

ولتقدير $\sigma^2_{Y/x}$ سوف نستخدم مجموع مربعات انحرافات المفردات حول خط الانحدار التقديري وقسمة هذا المجموع على درجات الحرية وهي تساوى n - 2 . وذلك لأننا استخدمنا درجتي حرية في تقدير β_0 و β_1 بنموذج الانحدار ، وعليه فإن :

$$\hat{\sigma}^2_{Y/x} = \frac{SS\,Res.}{n-2} = \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n-2} = \frac{1}{n-2}\left[\sum_{i=1}^{n}(y_i - \bar{y})^2 - \hat{\beta}_1\, SS_{xx}\right] \qquad (10)$$

أو

$$\hat{\sigma}^2_{Y/x} = \frac{1}{n-2}\left[SS_{yy} - \hat{\beta}_1\, SS_{xy}\right] \qquad (11)$$

حيث

$$SS_{yy} = \sum_{i=1}^{n}(y_i - \bar{y})^2 = \sum_{i=1}^{n} y_i^2 - \frac{\left(\sum_{i=1}^{n} y_i\right)^2}{n}$$

ومن المثال السابق نجد أن

$$SS_{yy} = 50075 - \frac{(767)^2}{12} = 1050.917$$

$$\hat{\sigma}^2_{Y/x} = \frac{1}{12-2}\left[1050.917 - (0.899)(964.583)\right] = 18.3757$$

لاحظ أن $\hat{\sigma}^2_{Y/x}$ يطلق عليها تسمية متوسط مربع الخطأ (mean square error) لنموذج خط الانحدار البسيط وهو يقيس متوسط مربعات انحرافات قيم العينة من خط الانحدار. ألاحظ أنه لرسم خط الانحدار التقديري يتطلب الأمر تحديد نقطتين فقط ثم رسم خط يمر من خلال هاتين النقطتين ، النقطة الأولى هي $(\bar{x}, \bar{y})$ وذلك لأن خط الانحدار يمر خلال تلك النقطة :

$$\hat{y} = \bar{y} + \hat{\beta}_1 (x - \bar{x})$$

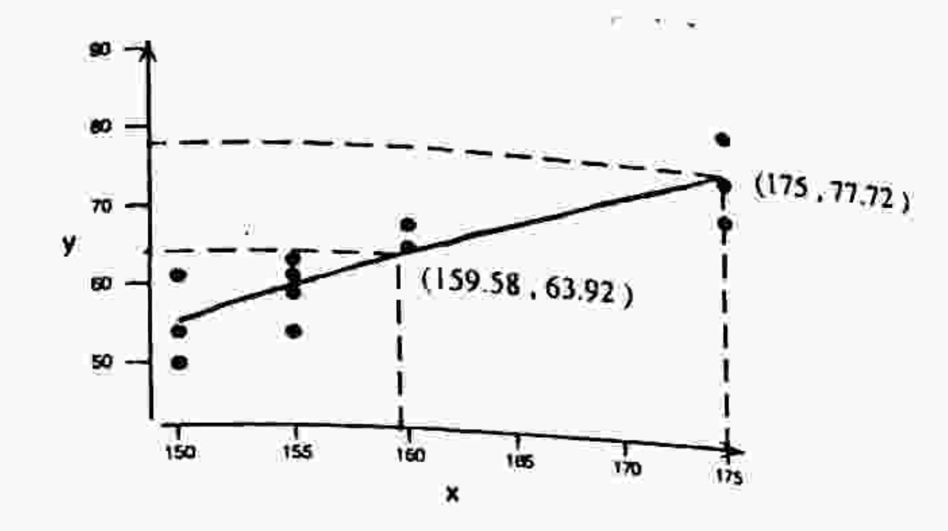
وبالتالي عندما $x = \bar{x}$ نجد أن

$$\hat{y} = \bar{y} + \hat{\beta}_1 (\bar{x} - \bar{x}) = \bar{y}$$

أما النقطة الثانية فهي $(x_0, \hat{y}_{x_0})$ حيث x_0 أبعد قيمة من قيم X عن $\bar{x}$ ، ثم أوجد :

$$\hat{y}_{x_0} = \hat{\beta}_0 + \hat{\beta}_1 x_0$$

ثم ارسم هذه النقطة ، وأخيراً فإن خط الانحدار التقديري يمر خلال النقطة $(0, \hat{\beta}_0)$.



شكل (7) :خط الانحدار التقديري .

3-9 الاستدلال الإحصائي للانحدار الخطى البسيط :

حيث أننا افترضنا أن نموذج الانحدار هو نموذج خطى بسيط وبالتالي نود تقييم جودة هذا النموذج، أى مدى توفيق معادلة الانحدار للبيانات ، وعادةً ما يتم تقييم نموذج الانحدار الخطى البسيط من خلال اختبار الفرضيات الإحصائية ذات العلاقة بمعلمات النموذج وتكوين فترات ثقة حول هذه المعلمات ، ولكن للقيام بمثل هذا الأجراء يتطلب الأمر معرفة خواص مقدرات المربعات الصغرى لمعلمات نموذج الانحدار وتوزيعاتها الاحتمالية وذلك لأنها المدخل الرئيسي لهذا التقييم .

حيث أنه عند دراستنا لأزواج النقاط (x_i , y_i) اعتبرنا y_i تمثل القيمة المشاهدة للمتغير العشوائي Y_i، ومن خلال النظر للصيغ الرياضية الخاصة بكل من $\hat{\beta}_0$ و $\hat{\beta}_1$ نلاحظ أن هاتين الصيغتين تعتمدان على قيم y_i، وعليه فإنهما يمثلان قيم مشاهدة لمتغيرات عشوائية وبالتالي يمكن إيجاد القيمة المتوقعة والتباين لهذه الصيغ إلا أننا سوف لن نتعرض هنا للكيفية التي يتم بها إيجاد كل من التوقع والتباين وذلك لأن ما يهمنا هنا هو الجانب التطبيقي وليس النظري ،ول لهذه المقدرات الخواص الآتية :

1- $\hat{\beta}_0$ و $\hat{\beta}_1$ مقدرين غير متحيزين أى أن

$$E(\hat{\beta}_1)=\beta_1 \quad و \quad E(\hat{\beta}_0)=\beta_0$$

-2 حيث اننا افترضنا بأن التوزيع الاحتمالي للخطأ توزيع طبيعي بمتوسط يساوى صفر وتباين يساوى σ^2 ، اى أن $\in_i \sim N(0,\sigma^2)\ , i = 1,2,\dots,n$ ، وعليه فإن توزيع المعاينة (التوزيع الاحتمالي) لكل من $\hat{\beta}_0$ و $\hat{\beta}_1$ يتبع التوزيع الطبيعي ويمكن الإثبات بأن

$$\hat{\beta}_0 \sim N(\beta_0 ,\ \frac{\sigma^2_{Y/x} \cdot \sum_{i=1}^{n} x_i^2}{n\sum_{i=1}^{n}(x_i - \bar{x})^2}) \qquad (12)$$

وأن

$$\hat{\beta}_1 \sim N(\beta_1 ,\ \frac{\sigma^2_{Y/x}}{\sum_{i=1}^{n}(x_i - \bar{x})^2}) \qquad (13)$$

-3 حيث أنه غالباً ما تكون $\sigma^2_{Y/x}$ مجهولة وبالتالي يتم استبدالها بالقيمة التقديرية لها وهى $\hat{\sigma}^2_{Y/x}$، وبما أن توزيع المعاينة لمعلمتي نموذج الانحدار توزيع طبيعي وبالتالي يمكن أيضاً الإثبات بأن :

$$\frac{\hat{\beta}_0 - \beta_0}{\hat{\sigma}_{\hat{\beta}_0}} \sim t_{(n-2)} \qquad (14)$$

وأن

$$\frac{\hat{\beta}_1 - \beta_1}{\hat{\sigma}_{\hat{\beta}_1}} \sim t_{(n-2)} \ . \qquad (15)$$

إن مما سبق يمكننا القيام باختبار الفرضيات الإحصائية وتكوين فترات الثقة الخاصة بمعلمات نموذج الانحدار الخطى البسيط وذلك كما يلي :

1 - اختبارات الفرضيات الإحصائية :

إذا تحقق الشرط الخاص بالخطأ $\in_i$، اى أن $\in_i \sim N(0,\sigma^2_{Y/x})$ والذي يمكن التحقق منه باستخدام ما يطلق عليه تسمية تحليل البواقي(residual analysis) وهو خارج نطاق هذا الكتاب، فإنه يمكن اختبار الفرضيات الآتية :

الفرضية	الإحصاءة	القرار
أ- اختبار من طرفين : $H_0:\beta_1 = b$ $H_1:\beta_1 \neq b$	$t = \frac{\hat{\beta}_1 - b}{\hat{\sigma}_{\hat{\beta}_1}}$ حيث $\hat{\sigma}_{\hat{\beta}_1} = \sqrt{\frac{\hat{\sigma}^2_{Y/x}}{SS_{xx}}}$	نرفض H_0 عند مستوى المعنوية α إذا كانت : $t \geq t_{\frac{\alpha}{2},n-2}$ أو $t \leq -t_{\frac{\alpha}{2},n-2}$
ب- اختبار من طرف واحد : $H_0:\beta_1 = b$ $H_1:\beta_1 > b$	$t = \frac{\hat{\beta}_1 - b}{\hat{\sigma}_{\hat{\beta}_1}}$	نرفض H_0 عند مستوى المعنوية α إذا كانت : $t \geq t_{\alpha,n-2}$
جـ- اختبار من طرف واحد : $H_0:\beta_1 = b$ $H_1:\beta_1 < b$	$t = \frac{\hat{\beta}_1 - b}{\hat{\sigma}_{\hat{\beta}_1}}$	نرفض H_0 عند مستوى المعنوية α إذا كانت : $t \leq -t_{\alpha,n-2}$

وبالمثل عند الاختبار بالنسبة للمعلمة β_0 .

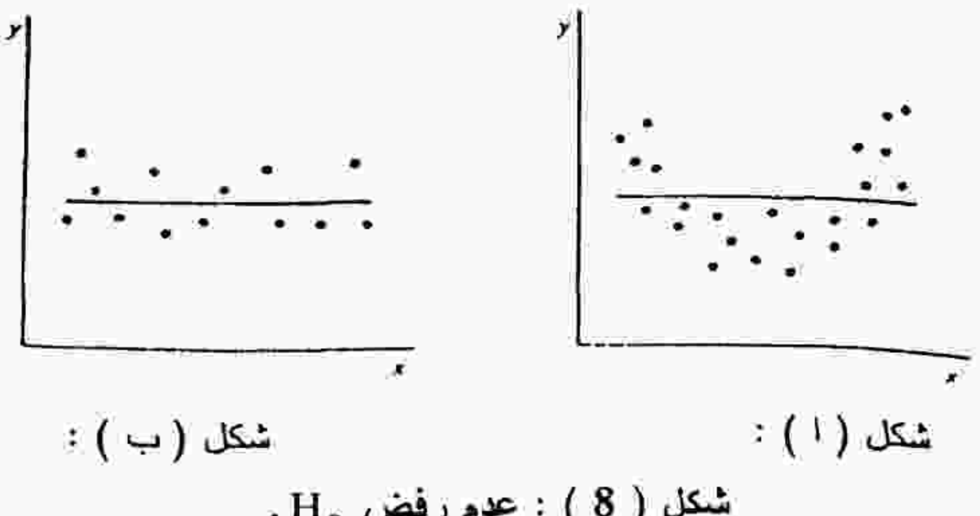
إن الحالة الخاصة والمهمة في الاختبارات أعلاه هي التى تكون فيهـا الفرضيـة علـى الصيغـة الآتية :

$$H_0:\beta_1 = 0$$
$$H_1:\beta_1 \neq 0$$

هذه الفرضية لها علاقة بمعنوية نموذج الانحدار وذلك لأنه ما تعنيه هذه الفرضية هو هل $\mu_{Y/x}$ تعتمد على x كما تم تحديدها بمعادلة الانحدار أم لا . وهذا يعنى أن الفرضية أعلاه يمكن إعادة كتابتها كما يلي :

$$H_0:\mu_{Y/x} = \beta_0$$
$$H_1:\mu_{Y/x} = \beta_0 + \beta_1\ x$$

وبالتالي عدم رفض H_0 يعنى أنه لا توجد علاقة خطية ما بين X و Y وهذا يؤدى بنا للقول إما أن X لها دور بسيط في تفسير التغير في Y كما في شكل (8-أ) وبالتالي فإن أفضل مقدر بالنسبة إلى $\mu_{Y/x}$ هو $\hat{y} = \bar{y}$ أو أنه في الحقيقة العلاقة ما بين X و Y ليست خطية كما في شكل (8 - ب) :



شكل (ب) :   شكل (ا) :

شكل (8) : عدم رفض H_0 .

وفي المقابل عند رفض H_0 فإن ذلك يعنى أن X تفسر التغير في Y ، وهذا يؤدى بنا للقول بأن النموذج الخطى هو النموذج المناسب ، أو أنه بالرغم من وجود تأثير خطى ولكن من الممكن الحصول على نتائج أفضل عند استخدام نموذج انحدار ذو رتبة أعلى . ومن بيانات المثال السابق يمكن اختبار الفرضية الآتية :

$$H_0 : \beta_1 = 0$$
$$H_1 : \beta_1 \neq 0$$

وذلك كما يلي : حيث أن

$$\hat{\sigma}^2_{\hat{\beta}_1} = \frac{\hat{\sigma}^2_{Y/x}}{SS_{xx}} = \frac{183.757}{1072.917} = 0.01713$$
$$\Rightarrow \quad \hat{\sigma}_{\hat{\beta}_1} = \sqrt{0.01713} = 0.1309$$

وعليه فإن

$$t = \frac{\hat{\beta}_1 - b}{\hat{\sigma}_{\hat{\beta}_1}} = \frac{0.899 - 0}{0.1309} = \frac{0.899}{0.1309} = 6.868$$

ومن جدول t وبدرجات حرية تساوى 10 و $\alpha = 0.05$ نجد أن

$$t_{\frac{\alpha}{2}, n-2} = t_{0.025, 10} = 2.228$$

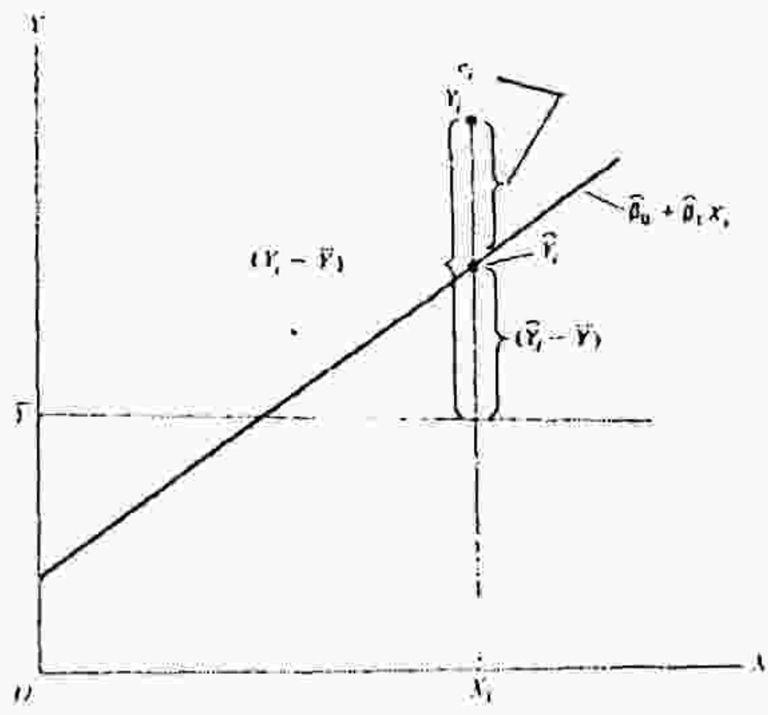
وحيث أن قيمة t المحسوبة وهي 6.868 أكبر من قيمة t الجدولية وهي 2.228 وبالتالي نرفض H_0 أي أن $\mu_{Y/x}$ تعتمد على x .

إن اختبار الفرضية $H_0:\beta_1 = 0$ أو المكافئة لها وهي $H_0:\mu_{Y/x} = \beta_0$ يمكن إجراءها بطريقة أخرى تعتمد على تجزئة أجمالي مجموع المربعات المصحح بالنسبة إلى y أي $\sum_{i=1}^{n}(y_i - \bar{y})^2$ وذلك كما يلي :-

$$\sum_{i=1}^{n}(y_i - \bar{y})^2 = \sum_{i=1}^{n}(y_i - \hat{y}_i + \hat{y}_i - \bar{y})^2$$

$$= \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 + \sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2 \qquad (16)$$

حيث $(y_i - \hat{y}_i)^2$ يمثل التغير غير المفسر آي الذي لا يعود إلى التغير في المتغير المستقل ، بينما $(\hat{y}_i - \bar{y})^2$ يمثل التغير المفسر آي الذي يعود إلى التغير في المتغير المستقل ويمكن توضيح ذلك بيانياً كما في الشكل الآتي :



شكل (9) : تجزئة أجمالي مجموع المربعات .

ويطلق على $\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$ تسمية مجموع مربعات البواقي ويرمز له بالرمز (SSRes) ، بينما يطلق على $\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2$ تسمية مجموع مربعات الانحدار ويرمز له بالرمز (SSReg) .
ألحظ أن

$$\sum_{i=1}^{n}(\hat{y}_i-\bar{y})^2=\sum_{i=1}^{n}(\hat{\beta}_0+\hat{\beta}_1 x_i-\bar{y})^2$$

$$=\sum_{i=1}^{n}(\bar{y}-\hat{\beta}_1\bar{x}+\hat{\beta}_1 x_i-\bar{y})^2$$

$$=\hat{\beta}_1^2\sum_{i=1}^{n}(x_i-\bar{x})^2=\hat{\beta}_1^2 SS_{xx} \quad (17)$$

وعليه فإن إجمالي مجموع المربعات المصحح والذي يرمز له بالرمز SST أمكن كتابته على النحو الآتي :

$$SST = SSReg. + SSRes. \quad (18)$$

درجات الحرية : $n - 1 = 1 + (n - 2)$

حيث SST له n-1 درجات حرية بينما .SSReg له درجة حرية واحدة ، وذلك لأن النموذج يتضمن متغير مستقل واحد فقط بينما .SSRes له " n-2 " درجات حرية وذلك لأن هناك معلمتين، ويمكن الإثبات بأن

$$E\left(\frac{SS\,Res.}{n-2}\right)=\sigma^2_{Y/x} \quad (19)$$

وأن

$$E(SS\,Reg.)=\sigma^2_{Y/x}+\beta_1^2\,SS_{xx} \quad (20)$$

وأن .SSRes و .SSReg مستقلين أيضاً ، وعليه إذا كانت الفرضية

$H_0:\beta_1=0$ أي أن $H_0:\mu_{Y/x}=\beta_0$ صحيحة فإن التوزيع الاحتمالي للإحصاءة

$$F_0=\frac{SSReg./1}{SSRes./n-2}=\frac{MS\,Reg.}{MSRES.}=\frac{MS\,Reg.}{\hat{\sigma}^2_{Y/x}} \quad (21)$$

يتبع توزيع f بدرجات حرية تساوي واحد و n-2 أي أن $F_0 \sim f_{1,n-2}$. وسوف نرفض H_0 إذا كانت $F_0 \geq f_{\alpha,1,n-2}$.

وبالتالي إذا كان H_0 صحيحاً فإن التناسب $\frac{MS\,Reg}{\hat{\sigma}^2_{Y/x}}$ سيكون قريب جداً من الواحد الصحيح ، أما إذا كان H_0 خطأ فإن هذا التناسب سيكون أكبر من الواحد ، وعادةً ما يتم وضع تحليل مجموع المربعات في جدول يطلق عليه تسمية جدول تحليل التباين (ANOVA) وذلك كما يلي:

المصدر Source	درجات الحرية d.f.	مجموع المربعات	متوسط المربعات	F_0
الإنحدار (Reg.)	1	$\hat{\beta}_1^2 SS_{xx}$	MSReg.	$\frac{MS\,Reg.}{MS\,Res.}$
البواقي (Res.)	n-2	$SS_{yy} - \hat{\beta}_1 SS_{xy}$	MSRes.	
المجموع (Tot.)	n-1	SS_{yy}		

فمن بيانات المثال السابق نجد أن :

$$SS\,Reg. = \hat{\beta}_1^2 SS_{xx} = (0.899)^2 (1072.917) = 867.133$$

$$SS\,Res. = 1050.917 - (0.899)(964.583) = 183.784$$

وعليه فإن جدول تحليل التباين يكون كالآتي :

مصدر الاختلاف	درجات الحرية	مجموع المربعات	متوسط المربعات	F_0
الإنحدار .Reg	1	867.133	867.133	$\frac{867.133}{18.378} = 47.183$
الخطأ .Res	10	183.784	18.378	
المجموع .Tot	11	1050.917		

ومن جدول f وبدرجات حرية واحد و 10 و $\alpha = 0.05$ نجد أن $f_{\alpha,1,n-2} = f_{0.05,1,10} = 4.96$
وحيث أن قيمة F المحسوبة وهي 47.183 أكبر من قيمة f الجدولية وهي 4.96 وعليه نرفض H_0 ، أي أن β_1 لا تساوي صفر ، وبعبارة أخرى $\mu_{Y/x}$ تعتمد على x ، وبالتالي يمكن استخدام العلاقة الخطية بينهما للتنبؤ بقيم y باستخدام قيم معينة من x .

II – التقدير ضمن فترة :

بالإضافة للتقدير بقيمة واحدة لكل من β_0 و β_1 و $\mu_{Y/x}$ واختبار الفرضيات الإحصائية الخاصة بها ، يمكن الحصول على تقدير ضمن فترة لكل منها ، وذلك لأن طول هذه الفترة يعتبر معيار لمدى ملائمة نموذج الانحدار في توفيق البيانات مدار البحث بصفة عامة .

أ- فترة ثقة لتقدير β_1 : $(1-\alpha)100\%$ فترة ثقة حول β_1 تكون كما يلي :

$$\hat{\beta}_1 - (t_{\frac{\alpha}{2},n-2})\,\hat{\sigma}_{\hat{\beta}_1} \le \beta_1 \le \hat{\beta}_1 + (t_{\frac{\alpha}{2},n-2})\,\hat{\sigma}_{\hat{\beta}_1} \qquad (22)$$

من المثال السابق نجد أنه إذا كانت $\alpha = 0.10$ فإن 90٪ فترة ثقة لتقدير β_1 تكون كالآتي :
من جدول t وبدرجات حرية تساوى 10 نجد أن $t_{\frac{\alpha}{2},n-2} = t_{0.05,10} = 1.812$ ، ومما سبق
نعلم أن $\hat{\sigma}_{\hat{\beta}_1} = 0.1309$ و $\hat{\beta}_1 = 0.899$ وعليه فإن

$$0.899 - (1.812)(0.1309) \le \beta_1 \le 0.899 + (1.812)(0.1309)$$
$$0.6618 \le \beta_1 \le 1.1362$$

وحيث أن هذه الفترة لا تحتوى على الصفر وعليه نرفض $H_0: \beta_1 = 0$.

بالمثل يمكن إيجاد فترة ثقة لتقدير β_0 : $(1-\alpha)100\%$ فترة ثقة حول β_0 تكون كالآتي :

$$\hat{\beta}_0 - (t_{\frac{\alpha}{2},n-2})\,\hat{\sigma}_{\hat{\beta}_0} \le \beta_0 \le \hat{\beta}_0 + (t_{\frac{\alpha}{2},n-2})\,\hat{\sigma}_{\hat{\beta}_0} \qquad (22)$$

ب- فترة ثقة لتقدير $\mu_{Y/x}$:

إن استخدامات نموذج الانحدار يمكن تقسيمها إلى قسمين هما : استخدام النموذج في تقدير $\mu_{Y/x}$ عند قيمة معينة للمتغير المستقل X أو التنبؤ بقيمة Y عند قيمة معينة للمتغير المستقل X ، ففي الحالة الأولى نحاول تقدير $\mu_{Y/x}$ لعدد كبير من التجارب عند قيمة معينة من X أما في الحالة الثانية فإننا نحول التنبؤ بنتيجة تجربة واحدة عند قيمة معينة من X .

ففي حالة تقدير $\mu_{Y/x_a} = \beta_0 + \beta_1 x_a$ حيث x_a تمثل قيمة معينة بالنسبة إلى X التي نريد عندها التقدير ، فإن تقدير المربعات الصغرى لهذه الدالة هو

$$\hat{y}_{x_a} = \hat{\mu}_{Y/X=x_a} = \hat{\beta}_0 + \hat{\beta}_1 x_a = \bar{y} + \hat{\beta}_1 (x_a - \bar{x}) \qquad (23)$$

وإن الخطأ المعياري بالنسبة إلى $\hat{y}_{x_a}$ هو :

$$\sigma_{\hat{y}_{x_a}} = \sigma_{y/x}\sqrt{\frac{1}{n}+\frac{(x_a-\bar{x})^2}{SS_{xx}}} \qquad (24)$$

وإن القيمة التقديرية لهذا الخطأ هي

$$\hat{\sigma}_{\hat{y}_{x_a}} = \hat{\sigma}_{y/x}\sqrt{\frac{1}{n}+\frac{(x_a-\bar{x})^2}{SS_{xx}}} \qquad (25)$$

ومن هذه الصيغة يتضح أن $\hat{\sigma}_{\hat{y}_{x_a}}$ تكون كبيرة كلما كانت $|x_a-\bar{x}|$ كبيرة ، وتكون أصغر ما يمكن عندما $x_a=\bar{x}$.

ويمكن إيجاد $(1-\alpha)100\%$ فترة ثقة حول μ_{Y/x_a} وذلك كما يلي :

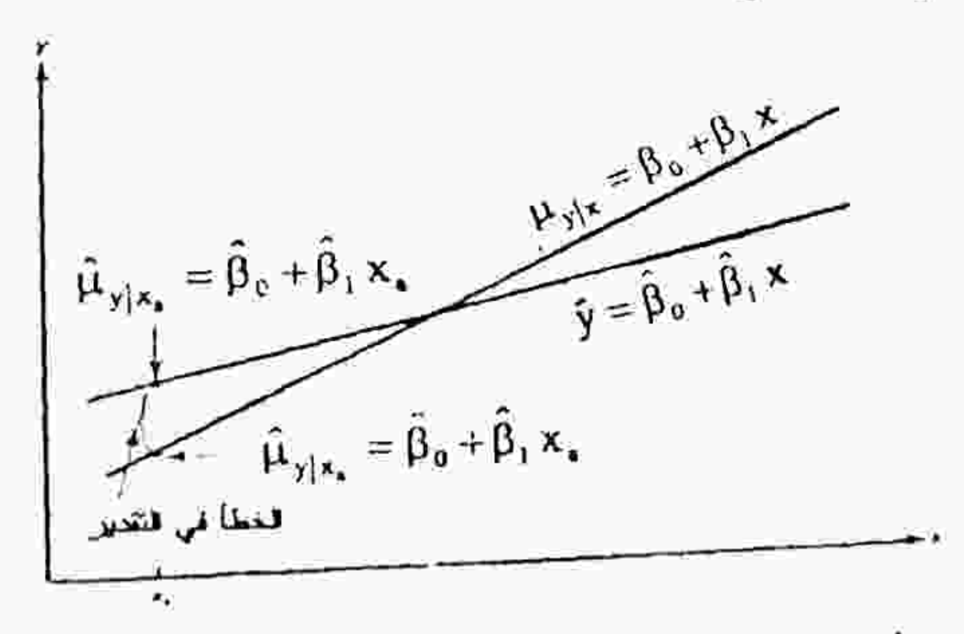
$$\hat{y}_{x_a} - (t_{\frac{\alpha}{2},n-2})\,\hat{\sigma}_{\hat{y}_{x_a}} \le \mu_{Y/x_a} \le \hat{y}_{x_a} + (t_{\frac{\alpha}{2},n-2})\,\hat{\sigma}_{\hat{y}_{x_a}} \qquad (26)$$

وعليه ستكون فترة الثقة أصغر ما يمكن عندما تكون $x_a=\bar{x}$ وستكون أكثر طولاً كلما زادت قيمة $|x_a-\bar{x}|$.

ملحوظة :

1- عندما تكون $x_a=0$ فإن فترة ثقة حول μ_{Y/x_a} هي نفس فترة الثقة حول β_0 .

2- إن الخطأ في تقدير $\mu_{Y/x}$ عند قيمة معينة بالنسبة إلى X هو المسافة العمودية ما بين خط المربعات الصغرى والخط الحقيقي للمتوسطات وسيكون هذا الخطأ أصغر ما يمكن عندما $x_a=\bar{x}$ كما في الشكل الآتي :



شكل (10) : الخطأ في تقدير $\mu_{Y/x}$ عندما $X = x_a$.

من المثال السابق يمكن إيجاد 95٪ فترة ثقة حول $\mu_{Y/x=x_a}$ وذلك كما يلي :

حيث أن $\hat{y} = -79.6 + 0.899\, x$ وعليه فإن $\hat{y}_{x_a} = -79.6 + 0.899\, x_a$

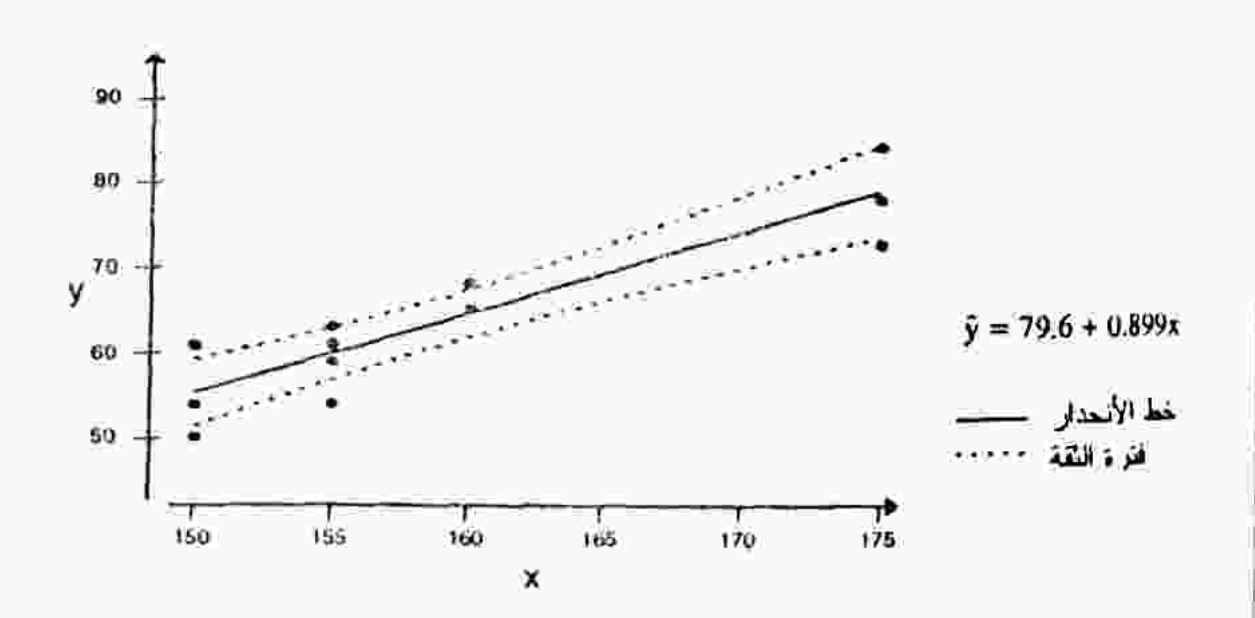
وإن القيمة التقديرية للخطأ المعياري تكون كالآتي :

$$\hat{\sigma}_{\hat{y}_{x_a}} = \left(\sqrt{18.3757}\right)\sqrt{\frac{1}{12} + \frac{(x_a - 159.58)^2}{1072.917}}$$

وإن $t_{\frac{\alpha}{2},n-2} = t_{0.025,10} = 2.228$ وعليه إذا كانت $x_a = 150, 155, 160, 175, \bar{x}$ فإن

x_a	$\hat{y}_{x_a} = -79.6 + 0.899\, x_a$	$\hat{\sigma}_{\hat{y}_{x_a}}$	$(t_{\frac{\alpha}{2},n-2})(\hat{\sigma}_{\hat{y}_{x_a}})$	الحد الأدنى	الحد الأعلى
150	55.25	1.762	3.926	51.324	59.176
155	59.75	1.375	3.064	56.686	62.814
160	64.24	1.239	2.760	61.48	67
175	77.73	2.367	5.274	72.456	83.004
159.58	63.86	1.237	2.756	61.104	66.616

ويمكن تمثيل ذلك بيانياً كما يلي :



شكل (11) : فترة ثقة حول μ_{Y/x_a} .

جـ - فترة التنبؤ : Prediction Interval

إن أحد التطبيقات المهمة لمنوذج الانحدار هو التنبؤ بقيمة جديدة أو مستقبلية للمتغير التابع Y عند قيمة معينة من قيم المتغير المستقل X ، إن هذه المفردة الجديدة مستقلة عن المفردات التي تم استخدامها في تكوين نموذج الانحدار، وبالتالي فإن فترة الثقة التي تعرضنا إليها سابقاً سوف لن تكون مناسبة، وذلك لأنها كانت حول متوسط التوزيع عندما $X = x_a$ ،علاوة على ذلك إنها كانت مبنية على أساس البيانات التي استخدمت في تكوين النموذج ، وعليه إذا كانت y_a تمثل المفردة المستقبلية عندما $X = x_a$ فإن $\hat{y}_{x_a} = \hat{\beta}_0 + \hat{\beta}_1 x_a$ تمثل المقدر لهذه المفردة والسؤال الذي نود الإجابة عليه الآن هو أين نتوقع أن تكون قيمة y_a الناتجة من التجربة عندما $X = x_a$ ؟ وللإجابة على هذا السؤال يجب تكوين فترة تنبؤ لها احتمال يساوى $1-\alpha$ بأن هذه الفترة تحتوى القيمة المستقبلية للمتغير التابع y عندما $X = x_a$.

إن $(1-\alpha)100\%$ فترة تنبؤ حول y_a تكون كالآتي :

$$\hat{y}_{x_a} - (t_{\frac{\alpha}{2},n-2})\hat{\sigma}_{\hat{y}_a} \le y_a \le \hat{y}_{x_a} + (t_{\frac{\alpha}{2},n-2})\hat{\sigma}_{\hat{y}_a} \qquad (27)$$

حيث

$$\hat{\sigma}_{\hat{y}_a} = \hat{\sigma}_{Y/x}\sqrt{1 + \frac{1}{n} + \frac{(x_a - \bar{x})^2}{SS_{xx}}}$$

فمن المثال السابق يمكن إيجاد 95٪ فترة تنبؤ حول y_a عندما $x_a = 150, 170, 180$ وذلك كما يلي :

$$x_a = 150 \Rightarrow \hat{\sigma}_{\hat{y}_{150}} = \left(\sqrt{18.3757}\right)\sqrt{1 + \frac{1}{12} + \frac{(150 - 159.58)^2}{1072.917}} = 4.635$$

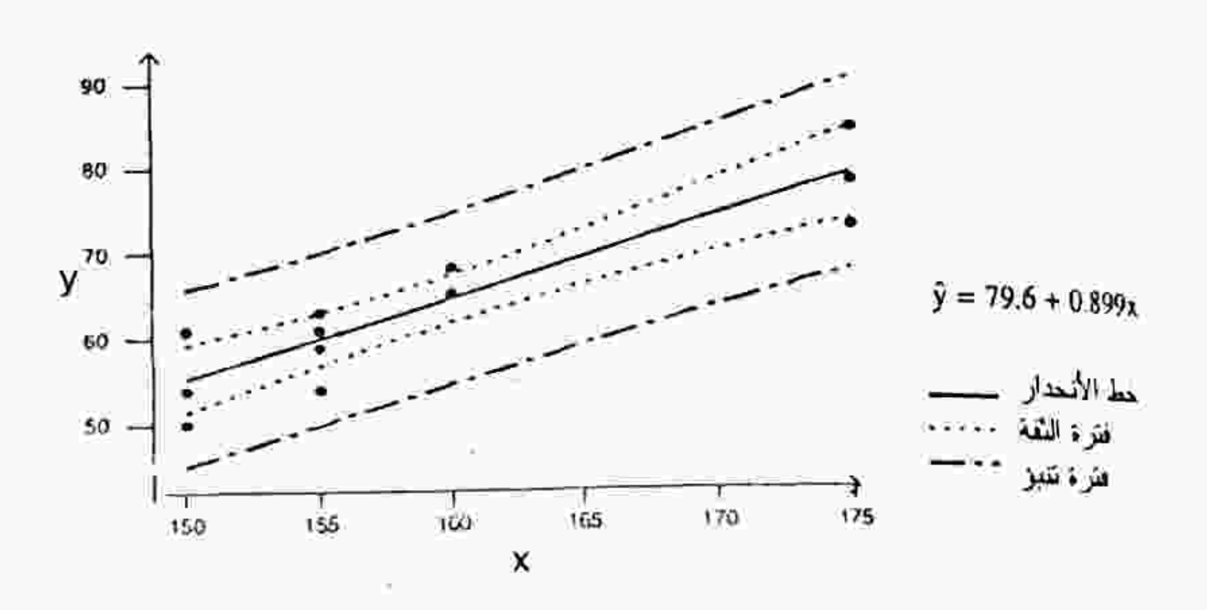
$$x_a = 170 \Rightarrow \hat{\sigma}_{\hat{y}_{170}} = \left(\sqrt{18.3757}\right)\sqrt{1 + \frac{1}{12} + \frac{(170 - 159.58)^2}{1072.917}} 4.666$$

$$x_a = 180 \Rightarrow \hat{\sigma}_{\hat{y}_{180}} = \left(\sqrt{18.3757}\right)\sqrt{1 + \frac{1}{12} + \frac{(180 - 159.58)^2}{1072.917}} = 5.201$$

وإن $t_{\frac{\alpha}{2},n-2} = t_{0.025,10} = 2.228$ وعليه فإن 95 ٪ فترة تنبؤ حول y_a عندما :

$$x_a = 150 \Rightarrow 55.25 - (2.228)(4.635) \le \hat{y}_{150} \le 55.25 + (2.228)(4.635)$$
$$44.923 \le \hat{y}_{150} \le 65.577$$
$$x_a = 170 \Rightarrow 73.23 - (2.228)(4.666) \le \hat{y}_{170} \le 73.23 + (2.228)(4.666)$$
$$62.834 \le \hat{y}_{170} \le 83.626$$
$$x_a = 180 \Rightarrow 82.22 - (2.228)(5.201) \le \hat{y}_{180} \le 82 + (2.228)(5.201)$$
$$70.632 \le \hat{y}_{180} \le 93.588$$

ويمكن تمثيل ذلك بيانياً كما يلي :



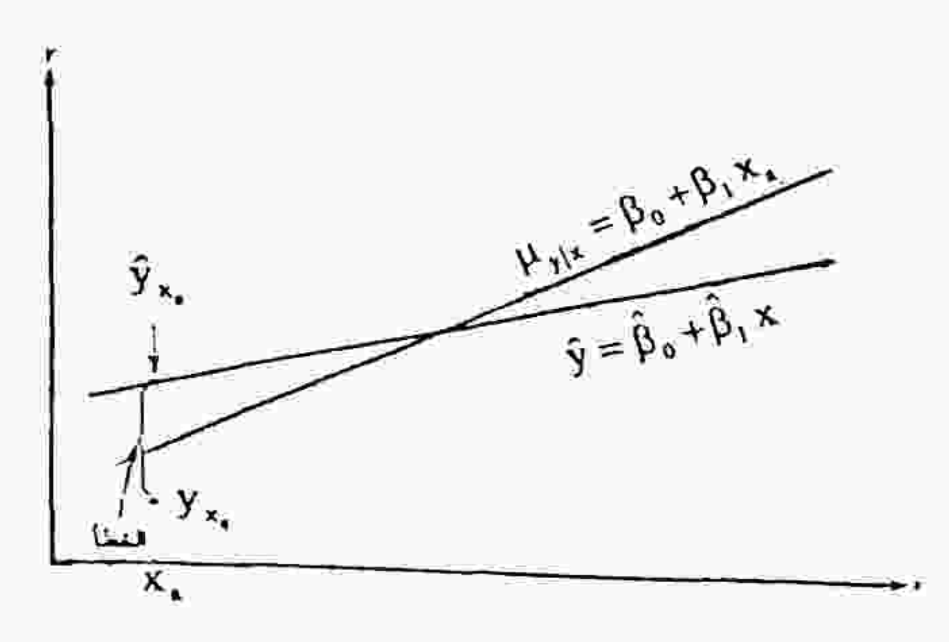
شكل (12) : فترة تنبؤ بمفردة مستقبلية y_a .

ملحوظة :

1- إن فترة التنبؤ بقيمة مستقبلية للمتغير التابع y ستكون أصغر ما يمكن عندما تكون $x_a = \bar{x}$ ويزداد طول هذه الفترة كلما زادت قيمة $|x_a - \bar{x}|$.

2- دائماً فترة التنبؤ تكون أطول من فترة الثقة حول $\mu_{y/x}$ عندما $x = x_a$ وذلك لأن هذه الفترة تعتمد على الخطأ الناتج من تقدير النموذج والخطأ المصاحب للمفردة المستقبلية ، أي أن :

$$\hat{\sigma}^2_{y_a} = \hat{\sigma}^2_{y/x} + \hat{\sigma}^2_{\hat{y}_a} \qquad (28)$$



شكل (13) : الخطأ في تقدير قيمة مستقبلية للمتغير التابع Y عندما $X = x_a$.

3- يجب ألا يستخدم نموذج الانحدار في تقدير $\mu_{Y/x}$ أو التنبؤ بقيمة مستقبلية للمتغير التابع Y لقيم معينة من X تقع خارج مدى قيم X التي ببيانات العينة ، وذلك لأن النموذج من الممكن أن يوفق البيانات بشكل جيد داخل مدى قيم X ولكنه لا يوفقها بشكل مقبول خارج مدى هذه القيم ، وبالتالي عدم محاولة تفادى مثل هذا الأمر قد يؤدى إلى خطأ في التقدير والتنبؤ وتكون هذه الأخطاء أكبر مما نتوقع .

4- في بعض الأحيان يكون لدى الباحث شعور بأن النموذج المناسب هو $y = \beta x + \in$ وهذا يعنى أن $y = 0$ عندما $x = 0$ ، إن هذا الافتراض يعتبر قوي جداً وليس هناك ما يبرره في معظم الأحيان حتى عند دراسة العلاقة ما بين الوزن والطول بالرغم من أن هذا النموذج مناسب ولكن سوف نتحصل على توفيق أفضل إذا تضمن النموذج β_0 وذلك لمحدودية مدى البيانات للمتغير المستقل .

لقد تعرضنا فيما سبق للمفاهيم الأساسية الخاصة بتوفيق نموذج الانحدار الخطى البسيط ، وذلك من خلال دراستنا لمثال الوزن والطول وإن خلاصة الحديث عن توفيق هذا النموذج بصفة عامة يمكن تلخيصه في النقاط التالية :

أ - وضع الصيغة الافتراضية لنموذج الانحدار .

ب - استخدام بيانات العينة لتقدير المعلمات المجهولة بذلك النموذج .

جـ - تحديد التوزيع الاحتمالي للخطأ العشوائي ثم تقدير معلمات التوزيع المجهولة .

د - اختبار مدى توفيق النموذج الافتراضي للبيانات قيد الدراسة .

هـ - إذا كان النموذج يوفق البيانات بالشكل المطلوب فإن هذا النموذج يمكن استخدامه في التقدير والتنبؤ .... الخ .

مثال (2) : بفرض أن شركة ليبيا للتأمين تحاول ربط حجم الخسائر الناتجة عن الحرائق بالمسافة ما بين مكان الحريق وأقرب محطة إطفاء للحريق من ذلك المكان ، ولهذا السبب قامت بجمع بيانات عن عينة تتكون من 15 حريق شبت مؤخراً في إحدى المدن فكانت النتائج كما يلي:

x_i :	5.44 2.88 7.36 3.68 4.96 8.8 1.12 4.8 4.16 6.88 3.36 1.76 9.76 7.68 6.08
y_i :	7.89 5.36 9.43 6.96 8.28 10.84 4.25 6.72 5.90 9.43 7.23 5.21 13.01 10.96 7.86

حيث x تمثل المسافة و y حجم الخسائر .

والمطلوب :

ا - رسم الشكل الانتشار واستنتاج النموذج المناسب لتوفيق وتحليل هذه البيانات من خلاله .

ب - إيجاد معادلة الانحدار التقديرية التي تم التوصل إليها من خلال الشكل الأنتشاري وتمثيلها بيانياً .

جـ - حساب تباين الانحدار التقديري .

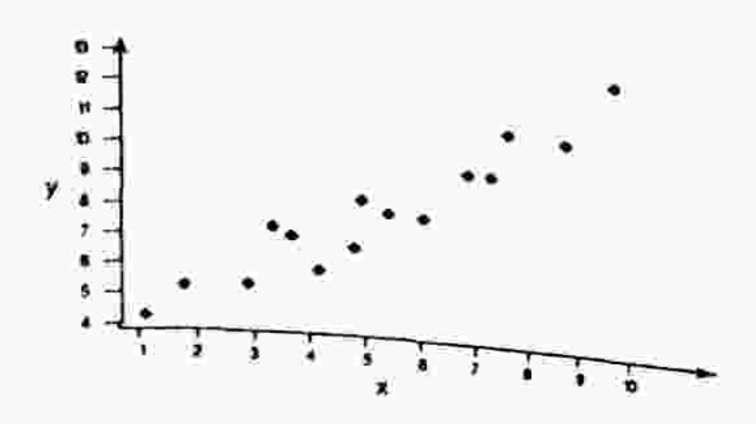
د - تكوين 95 ٪ فترة ثقة حول β_1 مع التعليق على النتائج .

هـ - إذا كانت المسافة ما بين مكان الحريق وأقرب محطة لإطفائه تساوى 5.6 كم فما هو حجم الخسائر المتوقع .

ر - تكوين 95 ٪ فترة تنبؤ حول $\hat{y}_{5.6}$.

الحل :

ا - إن الخطوة الأولى في تحليل هذه البيانات وكما أشرنا سلفاً هو رسم الشكل الأنتشاري للبيانات كما في الشكل الآتي :



شكل (14) : الشكل الأنتشارى للبيانات .

من خلال الرسم يتضح أن هناك علاقة خطية ما بين x و y وبالتالي فإن النموذج المناسب لهـ
هذه البيانات هو نموذج الانحدار الخطى البسيط والذي له الصيغة التالية :

$$y = \beta_0 + \beta_1 x + \in = \mu_{Y/x} + \in$$

ب - لإيجاد معادلة الانحدار الخطى البسيط التقديرية يتطلب الأمر حساب القيم التقديرية لمعم
الانحدار ، و سوف نوجد أولاً الحسابات الضرورية لذلك كما هو مبين بالجدول التالي :

x_i	y_i	x_i^2	y_i^2	$x_i y_i$
5.44	7.89	29.5936	62.252	42.922
2.88	5.36	8.2944	28.730	15.437
7.36	9.43	54.1696	88.925	69.405
3.68	6.96	13.5424	48.925	25.613
4.96	8.28	24.6016	68.558	41.069
8.8	10.84	77.4400	117.506	95.392
1.12	4.25	1.2544	18.063	4.760
4.8	6.72	23.0400	45.158	32.256
4.16	5.90	17.3056	34.810	24.544
6.88	9.43	47.3344	88.925	64.878
3.36	7.23	11.2896	52.273	24.293
1.76	5.21	3.0976	27.144	9.170
9.76	13.01	95.2576	169.260	126.978
7.68	10.96	858.9824	120.122	84.173
6.08	7.86	36.9664	61.780	47.789

ومن هذا الجدول نجد أن :

$$\sum_{i=1}^{15} y_i = 119.33 \text{ و } \sum_{i=1}^{15} x_i = 78.72 \text{ و } \sum_{i=1}^{15} y_i^2 = 1031.9 \text{ و } \sum_{i=1}^{15} x_i^2 = 502.17 \text{ و}$$

$$\sum_{i=1}^{15} x_i\, y_i = 708.68 \text{ و } \bar{y} = \frac{119.33}{15} = 7.955 \text{ و } \bar{x} = \frac{78.72}{14} = 5.248 \; .$$

وعليه فإن

$$SS_{xx} = \sum_{i=1}^{15} x_i^2 - \frac{\left(\sum_{i=1}^{15} x_i\right)^2}{n} = 502.17 - \frac{(78.72)^2}{15} = 89.05$$

$$SS_{yy} = \sum_{i=1}^{15} y_i^2 - \frac{\left(\sum_{i=1}^{n} y_i\right)^2}{n} = 1031.9 - \frac{(119.33)^2}{15} = 82.59$$

$$SS_{xy} = \sum_{i=1}^{15} x_i y_i - \frac{\left(\sum_{i=1}^{15} x_i\right)\left(\sum_{i=1}^{15} y_i\right)}{n} = 708.68 - \frac{(119.33)(78.72)}{15} = 82.44$$

وبالتالي فإن

$$\hat{\beta}_1 = \frac{SS_{xy}}{SS_{xx}} = \frac{82.44}{89.05} = 0.926$$

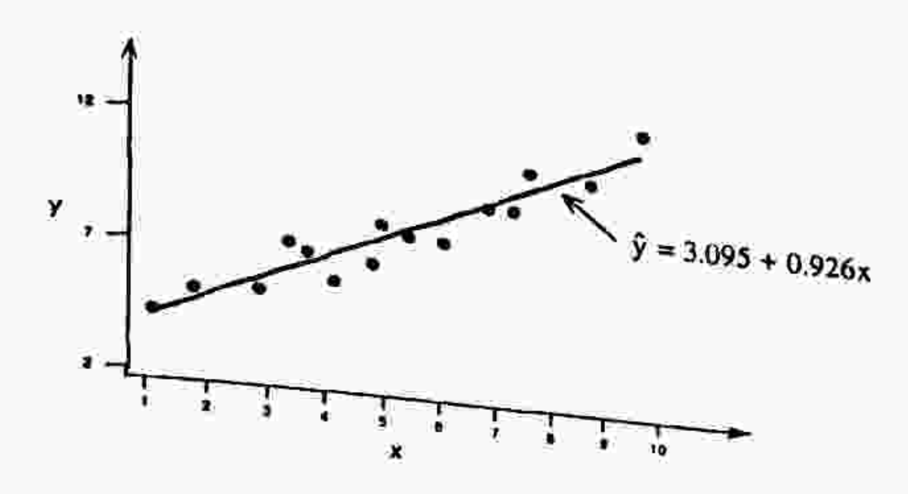
$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} = 7.955 - (0.926)(5.248)$$

$$= 3.095$$

وبذلك تكون معادلة الانحدار التقديرية لهذا النموذج كالآتي :

$$\hat{y} = 3.095 + 0.926\, x$$

وهي ممثلة بيانياً في الشكل الآتي :



شكل (15) : معادلة الانحدار التقديرية .

جـ - إن التوزيع الاحتمالي للخطأ العشوائي يتبع التوزيع الطبيعي بمتوسط يساوى صفر وتباين يساوى $\sigma^2_{Y/x}$.

إن متوسط التوزيع يساوى صفر يعنى أنه لكل قيمة ممكنة للمتغير المستقل X سيكون متوسط الأخطاء لسلسلة لانهائية من التجارب يساوى صفر والتباين ثابت لجميع قيم X ، بالإضافة إلى ذلك إن الخطأ المصاحب لأي قيمة معينة من قيم Y لا تأثير له على الأخطاء المصاحبة للقيم الأخرى .

إذن المعلمة الوحيدة المطلوب تقديرها لهذا التوزيع هي $\sigma^2_{Y/x}$ ويتم تقديرها كما يلي :

$$\hat{\sigma}^2_{Y/x} = \frac{SS\,Res.}{n-2} = \frac{1}{n-2}\left[SS_{yy} - \hat{\beta}_1 SS_{xy}\right]$$

$$= \frac{1}{15-2}\left[82.59 - (0.926)(82.44)\right]$$

$$= \frac{1}{13}\left[82.59 - 76.34\right]$$

$$= \frac{6.25}{13} = 0.4808$$

و...به فإن $\hat{\sigma}_{Y/x} = \sqrt{0.4808} = 0.6934$

د - إن اختبار جودة النموذج الافتراضي ، أى معرفة ما إذا كان المتغير المستقل يشارك بمعلومات تساعد في التنبؤ بقيم Y عند استخدام النموذج الخطى البسيط هو اختبار $H_0 : \beta_1 = 0$

وهذا يعني أنه لا توجد علاقة خطية ما بين حجم الخسائر والمسافة ما بين مكان الحريق وأقرب محطة لإطفائه .

إن هذه الفرضية مطابقة للفرضية $H_0: \mu_{Y/x} = \beta_0$ هذه الفرضية يمكن اختبارها كما أشرنا سابقاً ، إما باستخدام اختبار t أو باستخدام اختبار f وهذا الأخير يمكن حسابه من خلال تكوين جدول تحليل التباين ولتكوين هذا الجدول يتطلب الأمر حساب المقادير الآتية :

$$SS\,Reg = \hat{\beta}_1^2\, SS_{xx}$$
$$= (0.926)^2 (89.05) = 76.34$$
$$SST = SS_{yy} = 82.59$$

وعليه فإن جدول تحليل التباين يكون كما يلي :

المصدر .s.v	درجات الحرية d.f.	مجموع المربعات SS	متوسط المربعات MS	F_0
الإنحدار .Reg	1	76.34	76.34	$\frac{76.34}{0.4808} = 158.78$
البواقي .Res	13	6.25	0.4808	
المجموع .Tot	14	82.59		

وعليه من جدول f وبدرجات حرية و 1 و 13 و $\alpha = 0.05$ نجد أن $f_{0.05;1,13} = 4.67$ وحيث أن 4.67 أصغر من 158.78 ، وبالتالي نرفض H_0 أي أن العلاقة خطية ما بين حجم الخسائر والمسافة ما بين مكان الحريق وأقرب محطة لإطفائه .

هـ - كما سبق فإن 95 ٪ فترة ثقة لتقدير β_1 تكون كالآتي :

$$\hat{\beta}_1 - (t_{\frac{\alpha}{2}, n-2})\,\hat{\sigma}_{\hat{\beta}_1} \le \beta_1 \le \hat{\beta}_1 + (t_{\frac{\alpha}{2}, n-2})\,\hat{\sigma}_{\hat{\beta}_1}$$

وحيث أن $t_{\frac{\alpha}{2}, n-2} = t_{0.025,13} = 2.160$ وإن

$$\hat{\sigma}^2_{\hat{\beta}_1} = \frac{\hat{\sigma}^2_{Y/x}}{SS_{xx}} = \frac{0.4808}{89.05} = 0.0054 \;\Rightarrow\; \hat{\sigma}_{\hat{\beta}_1} = \sqrt{0.0054} = 0.0735$$

وعليه فإن

$$0.926 - (2.160)(0.0735) \leq \beta_1 \leq 0.926 + (2.160)(0.0735)$$
$$0.926 - 0.1588 \leq \beta_1 \leq 0.926 + 0.1588$$
$$0.7672 \leq \beta_1 \leq 1.0848$$

اى أن متوسط الزيادة في الخسائر يقدر ما بين 0.7672 و 1.0848 دينار كلما زادت المسافة بمقدار كيلو متر واحد .

و- حيث أن النتائج السابقة تشير إلى وجود علاقة بين حجم الخسائر والمسافة ما بين مكان الحريق وأقرب محطة لإطفائه ، وبالتالي يمكن استخدام هذه العلاقة للتنبؤ بحجم الخسائر ، فإذا كانت المسافة ما بين مكان الحريق وأقرب محطة لإطفائه تساوى 5.6 كم ، فإن حجم الخسائر المتوقع سيكون كالآتي :

حيث أن $x_a = 5.6$ وعليه فإن

$$\hat{y}_{x_a} = 3.095 + 0.926\ x_a$$
$$= 3.095 + (0.926)(5.6) = 8.2806$$

وإن

$$\hat{\sigma}_{\hat{y}_a} = \hat{\sigma}_{\hat{y}_{5.6}} = \hat{\sigma}_{Y/x}\sqrt{1 + \frac{1}{n} + \frac{(x_a - \bar{x})^2}{SS_{xx}}}$$
$$= (0.6934)\sqrt{1 + \frac{1}{15} + \frac{(5.6 - 5.248)^2}{89.05}} = 0.7166$$

إذن 95 ٪ فترة تنبؤ حول $\hat{y}_{5.6}$ تكون كالآتي :

$$\hat{y}_{x_a} - (t_{\frac{\alpha}{2},n-2})\hat{\sigma}_{\hat{y}_a} \leq \hat{y}_a \leq \hat{y}_{x_a} + (t_{\frac{\alpha}{2},n-2})\hat{\sigma}_{\hat{y}_a}$$
$$8.2806 - (2.160)(0.7166) \leq \hat{y}_{5.6} \leq 8.2806 + (2.160)(0.7166)$$
$$6.7327 \leq \hat{y}_{5.6} \leq 9.8285$$

هذا النموذج يقدر حجم الخسائر ما بين 6.7327 و 9.8285 دينار إذا كانت المسافة ما بين مكان الحريق و أقرب محطة لإطفائه تساوى 5.6 كم .

أخيراً يجب التنبيه إلى أنه وكما أشرنا سابقاً إلى عدم استخدام النموذج لتنبؤ خارج حدود قيم X ففي هذا المثال يجب ألا يستخدم النموذج في التنبؤ بقيم أقل من 1.12 أو أكبر من 9.70 كم ، وذلك لأنه من الممكن أن لا تكون العلاقة خطية ما بين $\mu_{Y/x}$ و X خارج حدود هذه القيم .

9-4 الانحدار المتعدد Multiple Regression

لقد أشرنا في البند السابق إلي أن تحليل الانحدار أسلوب يستخدم في تحديد نوعيـة العلاقـة مـا بين متغيرين أو أكثر ، وإن معظم التطبيقات العملية لتحليل الانحدار تستخدم نمـاذج أكثر تعقيدا مـن نموذج الانحدار الخطى البسيط . فمثلا عند دراستنا لمثال الأوزان والأطوال من الممكن أن يعتمد وزن الشخص على طوله ووزن أمه ووزن أبيه ، بالمثل عند دراسة الطلب علـى سـلعة معينـة لا يعتمد ذلك على سعر تلك السلعة فقط بل من الممكن أن يعتمد على أسعار سلع أخرى منافســة لهـا ودخل الفرد والحالة الاجتماعية ، أو عند دراسة سرعة الرياح من الممكن أن تعتمد على ارتفاعها ودرجة الحرارة والضغط الجوى ، وعليه يتطلب الأمر دراسة الانحدار المتعدد الذي يتضمن أكثر من متغير مستقل واحد . ففي هذا البند سوف نتناول مسألة تقدير أو التنبؤ بقيمة المتغير التـابع Y بناءاً على مجموعة من القياسات مأخوذة عن عدة متغيرات عشوائية مستقلة ولتكن :

$X_1, X_2, X_3, \dots, X_k$ وعليه إذا تم اختيـار عينـة عشوائية عددهـا n مـن المجتمـع الإحصـائي مدار البحث والتي يمكن تمثيلها بالمجموعـة $\{(x_{i1}, x_{i2}, \cdots, x_{ik}, y_i), i = 1,2,\cdots,n\}$ ، فإن نموذج الانحدار المتعدد يكون على الصورة الآتية :

$$y = \mu_{Y/\underline{x}} + \varepsilon \tag{29}$$

حيث

$$\mu_{Y/\underline{x}} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots \beta_k x_k$$

و $\underline{x} = (x_1, x_2, x_3, \dots, x_k)$

و β_0 تمثل قيمة $\mu_{Y/\underline{x}}$ عندما $x_1 = x_2 = \dots = x_k = 0$ بينمـا β_i و i =1,2,3,...,k تمثـل مـدى مشاركة المتغير المستقل x_i فـي هـذه العلاقـة . و β_i غير معلومـة وذلـك لأنهـا تمثـل معلمـات المجتمع . إن نموذج الانحدار المتعدد يطلق عليه تسـمية نمـوذج خطـى (linear model) طالمـا كانت المعلمات خطيـة حتى وإن تضمـن متغـيرات مسـتقلة غـير خطيـة , أي أنـه إذا كـانت $x_1 = x, x_2 = x^2, x_3 = x^3, \cdots, x_k = x^k$ فإن النموذج (29) سيكون على الصـورة الآتيـة :

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \cdots + \beta_k x^k + \varepsilon \tag{30}$$

فإذا افترضنا أن متوسط التوزيع الاحتمالي للخطأ ($\in$) يساوى صفر وبتبـاين يسـاوى σ^2 وعليـه فإن تباين Y يساوى $\sigma^2_{Y/\underline{x}}$ حيث $\sigma^2_{Y/\underline{x}} = \sigma^2$ وهو ثابت لأي مجموعة من المتغيرات المسـتقلة $x_1, x_2, \dots, x_k$ علاوة على ذلك إن الأخطاء مستقلة عن بعضها البعض .وكمـا أشـرنا فـي حالـة النموذج الخطى البسيط ، إن أول خطوة في تحليل الانحدار هي وضع الصيغة الافتراضية لنموذج

الانحدار ، فإذا افترضنا أن نموذج الانحدار كما هو موضح في المعادلة (29) فإن البيانات ستكون كما في الجدول الآتي :

y	x_1	x_2	$\cdots$	x_k
y_1	x_{11}	x_{12}	$\cdots$	x_{1k}
y_2	x_{21}	x_{22}	$\cdots$	x_{2k}
$\vdots$	$\vdots$	$\vdots$		$\vdots$
y_n	x_{n1}	x_{n2}	$\cdots$	x_{nk}

وبالتالي يمكن كتابة النموذج (29) كما يلي :

$$\begin{aligned} y_i &= \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \cdots + \beta_k x_{ik} + \epsilon_i \\ &= \beta_0 + \sum_{j=1}^{k} \beta_j x_{ij} \qquad , i = 1,2,3,\cdots,n \qquad (31) \end{aligned}$$

وحيث أن معلمات الانحدار $\beta_0, \beta_1, \ldots\ldots, \beta_k$ مجهولة ويجب تقديرها ، وبالتالي وكما في حالة نموذج الانحدار الخطى البسيط سوف نستخدم طريقة المربعات الصغرى (L.s.m) لتقدير هذه المعلمات وسيتم اختيار النموذج التقديري :

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \ldots + \hat{\beta}_k x_k \qquad (32)$$

بحيث يكون

$$SS\,Res. = \sum_{i=1}^{n} (y_i - \hat{\beta}_0 - \sum_{j=1}^{k} \hat{\beta}_j x_{ij})^2 \qquad (33)$$

أصغر ما يمكن .

إن مقدرات المربعات الصغرى لمعلمات النموذج يجب أن تفي :

$$\left.\frac{\partial SS\,Res.}{\partial \beta_0}\right|_{\hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_k} = -2 \sum_{i=1}^{n} (y_i - \hat{\beta}_0 - \sum_{j=1}^{k} \hat{\beta}_j x_{ij}) = 0$$

و

$$\left.\frac{\partial SS\,Res.}{\partial \beta_j}\right|_{\hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_k} = -2 \sum_{i=1}^{n} (y_i - \hat{\beta}_0 - \sum_{j=1}^{k} \hat{\beta}_j x_{ij}) x_{ij} = 0 \quad , j = 1,2,\ldots,k$$

وبتبسيط هذه المعادلات فإن المعادلات الطبيعية للمربعات الصغرى ستكون كالآتي :

$$
\begin{aligned}
&n\hat{\beta}_0+\hat{\beta}_1\sum_{i=1}^{n}x_{i1}+\hat{\beta}_2\sum_{i=1}^{n}x_{i2}+\ldots+\hat{\beta}_k\sum_{i=1}^{n}x_{ik}=\sum_{i=1}^{n}y_i\\
&\hat{\beta}_0\sum_{i=1}^{n}x_{i1}+\hat{\beta}_1\sum_{i=1}^{n}x_{i1}^2+\hat{\beta}_2\sum_{i=1}^{n}x_{i1}x_{i2}+\ldots+\hat{\beta}_k\sum_{i=1}^{n}x_{i1}x_{ik}=\sum_{i=1}^{n}x_{i1}y_i\\
&\quad\vdots\\
&\hat{\beta}_0\sum_{i=1}^{n}x_{ik}+\hat{\beta}_1\sum_{i=1}^{n}x_{ik}x_{i1}+\hat{\beta}_2\sum_{i=1}^{n}x_{ik}x_{i2}+\ldots+\hat{\beta}_k\sum_{i=1}^{n}x_{ik}^2=\sum_{i=1}^{n}x_{ik}y_i
\end{aligned}
\tag{34}
$$

وإنه من السهل حل هذه المعادلات إذا تمت كتابتها بطريقة المصفوفات .

إن أبسط نماذج الانحدار المتعدد هو ذلك النموذج الذي يحتوى على متغيرين مستقلين فقط وسوف نقتصر في دراستنا هنا على هذا النوع من النماذج الخطية في المعلمات فقط ، أى التي تكون على الصورة الآتية:

$$y=\beta_0+\beta_1x_1+\beta_2x_2+\varepsilon \tag{35}$$

حيث β_1 تقيس التغير فى $\mu_{Y/\underline{x}}$ $\left(\underline{x}=(x_1,x_2)\right)$ بكل وحدة تغير فى x_1 عندما تكون x_2 ثابتة و β_2 تقيس التغير فى $\mu_{Y/\underline{x}}$ بكل وحدة تغير فى x_2 عندما تكون x_1 ثابتة .

وفى هذه الحالة تكون المعادلات الطبيعية للمربعات الصغرى كما يلى :

$$
\begin{aligned}
&n\hat{\beta}_0+\hat{\beta}_1\sum_{i=1}^{n}x_{i1}+\hat{\beta}_2\sum_{i=1}^{n}x_{i2}=\sum_{i=1}^{n}y_i\\
&\hat{\beta}_0\sum_{i=1}^{n}x_{i1}+\hat{\beta}_1\sum_{i=1}^{n}x_{i1}^2+\hat{\beta}_2\sum_{i=1}^{n}x_{i1}x_{i2}=\sum_{i=1}^{n}x_{i1}y_i\\
&\hat{\beta}_0\sum_{i=1}^{n}x_{i2}+\hat{\beta}_1\sum_{i=1}^{n}x_{i1}x_{i2}+\hat{\beta}_2\sum_{i=1}^{n}x_{i2}^2=\sum_{i=1}^{n}x_{i2}y_i
\end{aligned}
\tag{36}
$$

وبحل منظومة هذه المعادلات نحصل على $\hat{\beta}_0$ و $\hat{\beta}_1$ و $\hat{\beta}_2$ ومهما يكون الأمر فإن الحسابات الضرورية ستكون طويلة وتحتاج إلى وقت وبالتالي يفضل استخدام الحاسب الآلي في مثل هذه الحالة .

مثال (3) : بفرض أن البيانات التالية تم جمعها لتحديد معادلة انحدار مناسبة تربط طول الطفل بعمره ووزنه عند الولادة :

الوزن عند الولادة x_2	العمر بالأيام (x_1)	طول الطفل بالسم (y)
2.75	78	57.5
2.15	69	52.8
4.41	77	61.3
5.52	88	67.0
3.21	67	53.5
4.32	80	62.7
2.31	74	56.2
4.30	79	68.5
3.71	102	69.2

ومن هذه البيانات أوجد : معادلة الانحدار التقديرية :

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2$$

والقيمة التقديرية للتنبؤ بطول طفل عمره 75 يوماً ووزنه 3.15 كجم عند الولادة .

الحـل :

إذن من البيانات أعلاه نجد أن

y_i	x_{i1}	x_{i2}	y_i^2	x_{i1}^2	x_{i2}^2	$y_i\, x_{i1}$	$y_i\, x_{i2}$	$x_{i1}\, x_{i2}$
57.5	78	2.75	3306.25	6084	7.5625	4485.0	158.125	214.50
52.8	69	2.15	2787.84	4761	4.6225	3643.2	113.520	148.35
61.3	77	4.41	3757.69	5929	19.4481	4720.1	270.333	339.57
67.0	88	5.52	4489.00	7744	30.4704	5896.0	369.840	485.76
53.5	67	3.21	2862.25	4489	10.3041	3584.5	171.735	215.07
62.7	80	4.32	3931.29	6400	18.6624	5016.0	270.864	345.60
56.2	74	2.31	3158.44	5476	5.3361	4158.8	129.822	170.94
68.5	94	4.30	4692.25	8836	18.4900	6439.0	294.550	404.20
69.2	102	3.71	4788.64	10404	13.7641	7058.4	256.732	378.42

ومن الجدول أعلاه نجد أن

$$\sum_{i=1}^{9} y_i = 548.7 \quad , \quad \sum_{i=1}^{9} x_{i1} = 729 \quad , \quad \sum_{i=1}^{9} x_{i2} = 32.68$$

$$\sum_{i=1}^{9} y_i^2 = 33774 \quad , \quad \sum_{i=1}^{9} x_{i1}^2 = 60123 \quad , \quad \sum_{i=1}^{9} x_{i2}^2 = 128.66 \ ,$$

$$\sum_{i=1}^{9} y_i x_{i1} = 45001 \ , \ \sum_{i=1}^{9} y_i x_{i2} = 2035.5 \ , \ \sum_{i=1}^{9} x_{i1} x_{i2} = 2702.4$$

وعليه فإن المعادلات الطبيعية ستكون كالآتي :

$$9\hat{\beta}_0 + 729\hat{\beta}_1 + 32.68\hat{\beta}_2 = 548.7$$

$$729\hat{\beta}_0 + 60123\hat{\beta}_1 + 2702.41\hat{\beta}_2 = 45001$$

$$32.68\hat{\beta}_0 + 2702.41\hat{\beta}_1 + 128.6602\hat{\beta}_2 = 2035.521$$

وبحل هذه المنظومة من المعادلات فإن تقديرات معلمات الانحدار ستكون كالآتي:

$$\hat{\beta}_0 = 20.108 \text{ و } \hat{\beta}_1 = 0.414 \text{ و } \hat{\beta}_2 = 2.035$$

وبالتالي فإن معادلة الانحدار التقديرية تكون كالآتي :

$$\hat{y} = 20.108 + 0.414\,x_1 + 2.035\,x_2$$

ومن هذا النموذج فإن الطول المتنبأ به لطفل عمره 75 يوما ووزنه عند الولادة 3.15 كجم هو

$$\hat{y} = 20.108 + (0.414)(75) + (2.035)(3.15) = 57.5$$

وكما أشرنا في حالة نموذج الانحدار الخطي البسيط ، إن الهدف الرئيسي من إيجاد معادلة الانحدار هو الحصول على مقياس يفيد في التنبؤ بقيم المتغير التابع عند قيم معروفة للمتغير المستقل ، وللوصول لهذا الهدف يجب أن نقيم أولاً مدى دقة معادلة الانحدار التقديرية في توفيق البيانات قيد التحليل ، وعند مناقشتنا لنموذج خط الانحدار البسيط أوضحنا أن الخطأ في التنبؤ بقيمة معينة للمتغير التابع(Y_i) رمزنا له بالرمز ε_i وعليه يمكن قياس مدى دقة تنبؤ معادلة الانحدار التقديرية من خلال تفحص قيم ε_i وإن تباين هذا الخطأ (متوسط مربع الانحرافات حول خط الانحدار) يستخدم كمقياس لانتشار قيم Y_i حول خط الانحدار . هذه القيمة رمزنا لها بالرمز $\sigma^2_{Y/x}$ وهو ببساطة عبارة عن تباين قيم Y_i حول خط الانحدار. بالمثل في الانحدار المتعدد هناك مقياس مشابه لهذا المقياس يقيس انتشار قيم Y_i حول فضاء الانحدار ، وفي حالة ما يكون هناك متغيرين مستقلين فإن هذا المقياس يرمز له بالرمز $\sigma^2_{Y/\underline{x}}$ ويتم تقديره كما يلي :

$$\hat{\sigma}^2_{Y/\underline{x}} = \frac{1}{n-3}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \qquad , \underline{x} = (x_1 , x_2) \qquad (37)$$

وبصفة عامة ، إن دقة التنبؤ لمعادلة الانحدار التقديرية التي تتضمن k من المتغيرات المستقلة

أى التي على الصورة $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + ... + \hat{\beta}_k x_k$

يمكن التعبير عنها كما يلي :

$$\hat{\sigma}^2_{Y/\underline{x}} = \frac{1}{n-k-1} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \quad , \quad \underline{x} = (x_1, x_2, x_3, ..., x_k) \qquad (38)$$

حيث k تمثل عدد المتغيرات المستقلة في النموذج .إذن من البيانات السابقة والتعويض عن قيم x_1, x_2 في معادلة الانحدار التقديرية نجد أن :

$\hat{y}_i$	$(y_i - \bar{y}_i)^2$	$(y_i - \hat{y}_i)^2$	$(\hat{y}_i - \bar{y})^2$
57.9745	12.0201	0.225	8.9551
53.0305	66.6999	0.0531	62.9880
60.9303	0.1109	0.1367	0.0013
67.7376	36.3971	0.5440	45.8410
54.3543	55.7561	0.7298	43.7278
61.9896	3.0073	0.5047	1.0457
55.4253	22.7243	0.6002	30.7104
67.7450	56.7461	0.5700	45.9412
69.8593	67.7823	0.4347	79.0731

ومن هذا الجدول نجد أن

$$SS\,Reg. = \sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2 = 317.46 \quad \text{و} \quad SS\,Res. = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = 3.78$$

وأن

$$SST = \sum_{i=1}^{n} (y_i - \bar{y})^2 = 321.24$$

ويمكن وضع هذه النتائج في جدول تحليل التباين وذلك كما يلي :

المصدر	درجات حرية d f	مجموع المربعات S S	متوسط المربعات ms	F المحسوبة
الانحدار Reg	2	317.46	158.73	$F_0 = 251.95$
الخطأ Res	6	3.78	0.63	
المجموع Tot	8	321.24		

وإن القيمة التقديرية لتباين الخطأ تكون كما يلي :

$$\hat{\sigma}^2_{Y/\underline{x}} = \frac{SS\,Res.}{n-3} = \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n-3} = \frac{3.78}{6} = 0.63$$

ومن جدول تحليل البيانات يمكن معرفة في ما إذا كان هناك علاقة بين أي من المتغيرين المستقلين والمتغير التابع وذلك باختبار الفرضية الآتية :

$$H_o : \beta_1 = \beta_2 = 0$$

H_1 : H_o ليس صحيحاً

وكما في حالة الانحدار الخطي البسيط إن هذه الفرضية يتم اختبارها باستخدام اختبار F وذلك على النحو الآتي :

$$F_0 = \frac{MS\,Reg.}{MS\,Res.} = \frac{158.73}{0.63} = 251.95$$

ومن جدول F وبدرجات حرية 2 و 6 و $\alpha = 0.05$ نجد أن $f_{0.05,2,6} = 5.14$ وحيث أن F_0 المحسوبة أكبر من 5.14 وعليه نرفض H_o ويمكن القول بأن هناك علاقة ما بين x_1 و x_2 أو كلاهما والمتغير التابع Y .

9 - 5 الارتبـــــاط Correlation

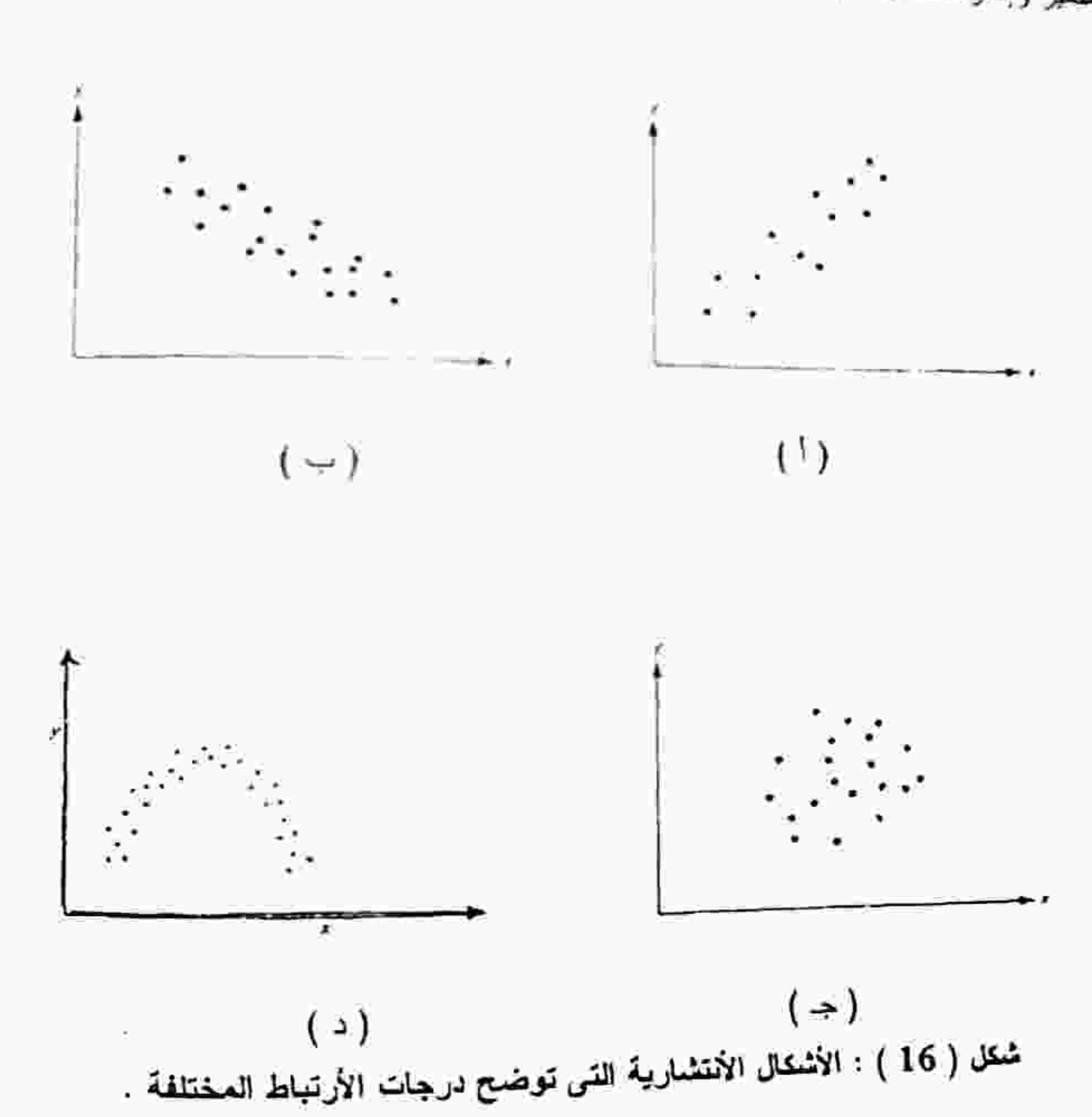
الارتباط هو الموضوع الذي يهتم ويبحث في العلاقة بين المتغيرات من حيث القوة والاتجاه ، وعليه فإن الارتباط معيار يقيس قوة واتجاه تلك العلاقة ويطلق عليه تسمية ارتباط بسيط إذا كان يبحث في قوة واتجاه العلاقة بين متغيرين فقط ، ومتعدد إذا كان يهتم بدراسة قوة واتجاه العلاقة بين عدة متغيرات ، وجزئي إذا كان يبحث في قوة واتجاه العلاقة بين مجموعة من المتغيرات مع عزل تأثير بعض المتغيرات الأخرى .

9-5-1 الارتباط الخطي البسيط Simple Linear Correlation

إن تحليل الانحدار يهتم بإيجاد علاقة بين متغيرين أو أكثر ثم استخدام هذه العلاقة في التقدير والتنبؤ ، بينما تحليل الارتباط يهتم بقياس درجة أو قوة العلاقة ما بين متغيرين أو أكثر . فمثلاً دراسة العلاقة ما بين التدخين (X) والإصابة بأمراض الرئة (Y) أو العلاقة ما بين معدل الجريمة (X) ومعدل البطالة (Y) ... الخ .إن تحليل الارتباط يهدف لقياس قوة مثل تلك

العلاقة والتعبير عنها بعدد واحد يطلق عليه تسمية معامل الارتباط ، ويعرف معامل الارتباط الخطي على أنه مقياس للعلاقة ما بين المتغيرين العشوائيين X و Y . فإذا كانت (x_i, y_i) حيث $i = 1, 2, \ldots, n$ تمثل عينة عشوائية ثنائية من القياسات وعددها n وتم رسم الشكل الانتشاري لهذه البيانات فإنه يمكننا التوصل إلى نوعية العلاقة ما بين المتغيرين X و Y :

أ- إذا كانت القيم الكبرى للمتغير X تقابلها قيم كبرى للمتغير Y وبالتالي القيم الصغرى للمتغير X تقابلها قيم صغرى للمتغير Y فإن العلاقة وطردية ما بين المتغيرين وإشارة معامل الارتباط موجبة كما في شكل (16- أ) . أما إذا كانت القيم الكبرى للمتغير X تقابلها قيم صغرى للمتغير Y وبالتالي القيم الصغرى للمتغير X تقابلها قيم كبرى للمتغير Y فإن العلاقة عكسية ما بين المتغير وإشارة معامل الارتباط سالبة كما في شكل (16 - ب) .



شكل (16) : الأشكال الأنتشارية التى توضح درجات الأرتباط المختلفة .

ب - إذا كانت قيم المتغير X يبدو أنها تتناظر مع قيم المتغير Y بشكل عشوائي فإن قيمة معامل الارتباط قريبة جداً من الصفر وتكون الخلاصة هي عدم وجود علاقة خطية ما بين المتغيرين كما في شكل (16 - جـ) .ألحظ أنه في بعض الأحيان تكون قيم معامل الارتباط تساوي صفر ولكن هذا لا يعنى عدم وجود علاقة ما بين المتغيرين وذلك لأنه من الممكن وجود علاقة ولكنها غير خطية كما في شكل (16 - د) . إن من أكثر المقاييس استخداماً لقياس قوة العلاقة ما بين متغيرين أو ظاهرتين هو "معامل بيرسون للارتباط" الذي يرمز له بالرمز r ويطلق عليه تسمية معامل ارتباط العينة ، وهو معرف كما يلي :

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\left[\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2\right]^{\frac{1}{2}}} \qquad (39)$$

وبقسمة كل من البسط والمقام على n والقيام ببعض العمليات الجبرية البسيطة يمكن إعادة كتابة المعادلة (39) كما يلي :

$$r = \frac{n\sum_{i=1}^{n} x_i\, y_i - \left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right)}{\left\{\left[n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2\right]\left[n\sum_{i=1}^{n} y_i^2 - \left(\sum_{i=1}^{n} y_i\right)^2\right]\right\}^{\frac{1}{2}}} \qquad (40)$$

إن هذا المقياس يمكن استخدامه لأي بيانات عددية بدون أى شروط تتعلق بوحدة قياس البيانات أو نوعية التوزيع ، بالرغم أنه من الصعب تفسيره إذا كانت وحدة القياس للبيانات أقل من القياس الفتروى (انظر فصل 11) وإنه يفي بالشروط المطلوبة لأي مقياس مقبول كمقياس للارتباط .

مثال (4) : اختيرت عينة عشوائية من 8 أشخاص من مجتمع كبير من الأشخاص ووجه لكل منهم سؤال عن عمر أبيه وعمر أمه عند ولادته فكانت النتائج كما يلي :

عمر الأم (x_i)	27	23	18	29	22	19	25	20
عمر الأب (y_i)	32	24	18	30	20	29	25	23

من هذه البيانات هل يمكن القول بوجود علاقة خطية ما بين عمر الأب وعمر الأم ؟

الحـل :

حيث أن

$$\sum_{i=1}^{8} y_i = 201 \quad , \quad \sum_{i=1}^{8} x_i^2 = 4293 \quad , \quad \sum_{i=1}^{8} x_i = 183$$

$$\sum_{i=1}^{8} x_i y_i = 4686 \quad , \quad \sum_{i=1}^{8} y_i^2 = 5219$$

وعليه فإن

$$r = \frac{8(4686)-(183)(201)}{\left\{\left[(8)(4293)-(183)^2\right]\left[8(5219)-(201)^2\right]\right\}^{\frac{1}{2}}}$$

$$= \frac{37488-36783}{\left\{[34344-33489][41752-40401]\right\}^{\frac{1}{2}}}$$

$$= \frac{705}{\left[(855)(1351)\right]^{\frac{1}{2}}} = \frac{705}{1074.758} = 0.656$$

ومن هذه القيمة يمكن القول بأنه هناك علاقة خطية ما بين عمر الأب وعمر الأم ولكنها ليست قوية .

ألحظ أن r^2 يطلق عليها تسمية معامل تحديد العينة (Sample coefficient of determination) (وهو يعبر عن نسبة الاختلاف الكلي في قيم المتغير Y التي يمكن تفسيرها من خلال علاقته الخطية مع قيم المتغير X ، وعليه عندما $r = 0.656$ فإن ذلك يعني أن 43٪ من إجمالي الاختلاف في قيم Y بهذه العينة يتم تفسيره بالعلاقة الخطية مع قيم X .

حيث أن معامل ارتباط العينة r يتم حسابه من بيانات عينة عشوائية ، وعليه لعينات مختلفة من نفس المجتمع الإحصائي مدار البحث ستكون له قيم مختلفة ، وبالتالي يمكن التفكير في r كتقدير بقيمة واحدة لمعامل الارتباط الخطي الحقيقي للمجتمع الإحصائي بأكمله والذي سنرمز له بالرمز ρ وعليه عندما تكون قيمة r قريبة من الصفر فإنه من الممكن القول بأن $\rho = 0$ ، ولكن إذا كانت

قيمة r قريبة من "+1" أو "-1" ، فإنها تشير إلى أن $\rho \neq 0$ ، وعليه نحن بحاجة إلى معرفة معنوية r ولاختبار هذه المعنوية يتطلب الأمر معرفة توزيع المعاينة لهذا المعامل .

وإذا افترضنا أن (X_i, Y_i) حيث $i=1,2,\cdots,n$ تشكل عينة عشوائية من مجتمع طبيعي ثنائي وكانت $\rho=0$ فإنه ثبت من الناحية النظرية بأن توزيع المعاينة (التوزيع الاحتمالي) للإحصاءه

$$T = r\sqrt{\frac{n-2}{1-r^2}} \qquad (41)$$

يتبع توزيع t وبدرجات حرية تساوي n -2 . وعليه يمكن استخدام هذه الإحصاءة كأساس لاختبار الفرضية $H_0 : \rho = 0$ مقابل جميع البدائل الممكنة ، واستخدام جداول توزيع t كما سبق في حالة الاختبار للمتوسط أو للفرق ما بين متوسطي مجتمعين من حيث قبول أو رفض الفرضية .

مثال (5) : استخدم بيانات المثال (4) لاختبار الفرضية التالية :

$$H_0 : \rho = 0$$
$$H_1 : \rho \neq 0$$

عند مستوى المعنوية $\alpha = 0.05$.

الحل :

حيث أن

$$t = r\sqrt{\frac{1-2}{1-r^2}} = (0.656)\sqrt{\frac{8-2}{1-(0.656)^2}} = 2.458$$

$$= (0.656)\sqrt{\frac{8}{057}} = 2.458$$

ومن جدول t وبدرجات حرية تساوي 6 نجد أن

$$t_{\frac{\alpha}{2}, n-2} = t_{0.025, 6} = 2.447$$

وحيث أن 2.458 أكبر من 2.447 ، وعليه لا توجد معلومات كافية لرفض H_0 .

حيث أن الطريقة التي سبق وأن تعرضنا إليها يمكن تطبيقها عندما يكون فرض العدم $\rho = 0$ فقط ، وبالتالي سوف نتعرض الآن إلى طريقة أخرى أكثر عمومية ، فعندما ρ لا تساوى صفر فإن توزيع المعاينة لمعامل ارتباط العينة (r) سيكون أكثر التواءً كلما اقتربت r من "-1" أو

$+1$ • ولاختبار أن $H_0: \rho = \rho_0$ حيث $\rho_0 \neq 0$ مقابل جميع البدائل الممكنة فإننا نستخدم

الاحصاءة :

$$Z = \frac{W - \mu}{\sigma}$$

(42) $$W = \frac{1}{2} \ln\left(\frac{1+r}{1-r}\right)$$ حيث Z تتبع التوزيع الطبيعي المعياري تقريباً وإن المتغير

(حيث ln ترمز للوغاريتم اساسه e) ، يتبع التوزيع الطبيعي تقريبياً بمتوسط يساوي $\mu = \frac{1}{2}\ln\left(\frac{1+\rho}{1-\rho}\right)$ وتباين يساوي $\sigma^2 = \frac{1}{n-3}$. والجدول (13) يعطي قيم w المناظرة إلى $r = 0.00 , 0.01 , 0.02 , 0.03 ,..., 0.99$. ألحظ أنه هذا الجدول يعطي القيم الموجبة فقط بالنسبة إلى r ، ولكن إذا كانت r سالبة فإننا ننظر إلى (r -) وأخذ القيمة السالبة المقابلة إلى w ، وسوف نرفض H_0 عند مستوى المعنوية α كما في حالة الاختبار للمتوسط أو لأي اختبار آخر يستخدم فيه التوزيع الطبيعي المعياري ، وإن التقريب أعلاه سيكون جيد إذا كانت $n \geq 25$.

مثال (6) : في المثال (5) أختبر الفرضية التالية :

$$H_0: \rho = 0.75$$
$$H_1: \rho \neq 0.75$$

عند مستوى المعنوية $\alpha = 0.05$.

الحل :

حيث أن $n = 8$ و $\rho = 0.75$ نجد أن

$$\sigma = \sqrt{\frac{1}{8-3}} = \sqrt{\frac{1}{5}} = 0.447 \text{ و } \mu = \frac{1}{2}\ln\left(\frac{1+0.75}{1-0.75}\right) = 0.973$$

وحيث أن $r = 0.656 \cong 0.66$ ، وعليه من جدول (13) نجد أن $w = 0.79281$ وبالتالي فإن :

$$z = \frac{0.79281 - 0.973}{0.447} = -0.403$$

ومن جدول التوزيع الطبيعي المعياري نجد أن $z_{\frac{\alpha}{2}} = z_{0.025} = 1.96$ ، وحيث أن $-0.403 > -1.96$ وعليه نرفض H_0 عند مستوى المعنوية 5 ٪ .

لاحظ أنه يمكن إيجاد $(1-\alpha)100\%$ فترة ثقة لتقدير ρ وذلك كما يلي :

1- التحويل من r إلى w باستخدام جدول (13) .

2- استخدام التوزيع الطبيعي لتكوين فترة ثقة حول μ وذلك كما يلي :

$$w - \frac{z_{\frac{\alpha}{2}}}{\sqrt{n-3}} \le \mu \le w + \frac{z_{\frac{\alpha}{2}}}{\sqrt{n-3}}$$

3- تحويل الناتج المتحصل عليه في الخطوة (2) إلى ρ باستخدام جدول (13) .

مثال (7) : إذا كانت $r = 0.62$ و $n = 30$ أوجد 95 ٪ فترة ثقة حول ρ .

الحل :

من جدول (13) نجد أنه عندما $r = 0.62$ تكون $w = 0.725$ ومن جدول التوزيع الطبيعي نجد أن $z_{\frac{\alpha}{2}} = z_{0.025} = 1.96$ وعليه فإن 95 ٪ فترة ثقة حول μ تكون كالآتي :

$$0.725 - \frac{1.96}{\sqrt{27}} \le \mu \le 0.725 + \frac{1.96}{\sqrt{27}}$$

أو

$$0.348 \le \mu \le 1.102$$

وبالنظر إلى قيم r التي قريبة جداً من $w = 0.348$ و $w = 1.102$ بجدول (13) نجد أن 95 ٪ فترة ثقة لتقدير ρ تكون كالآتي : $0.33 \le \rho \le 0.80$.

وبنهاية هذا البند نود أن نؤكد مرة أخرى إلى أنه إذا كانت العينة من توزيع طبيعي ثنائي فإن اختبار $\rho = 0$ يكافئ اختبار الاستقلالية ما بين المتغيرين أو الظاهرتين مدار البحث ، وإن أى قيمة لمعامل ارتباط المجتمع الإحصائي (ρ) بصرف النظر عن مدى قربها من " ± 1 " ما هي إلا مؤشر على قوة العلاقة ما بين المتغيرين ولكنها لا تفسر السبب في وجود هذه العلاقة ، وإن الأسلوب الذي اتبعناه في اختبار الفرضيات الخاصة بمعامل الارتباط تكون صحيحة فقط ، إذا كان الافتراض بان التوزيع توزيع طبيعي ثنائي تقريباً متحقق ، أما إذا كان هذا الافتراض غير متحقق فإنه يجب استخدام الأساليب اللامعلمية .

9-5-2 الارتباط المتعدد والارتباط الجزئي Multiple and partial correlation

إن مفهوم الارتباط الخطي البسيط ومعامل التحديد الذي سبق وأن تعرضنا إليه يُعد مقياس جيد لمدى جودة نموذج الانحدار الخطي البسيط في توفيق البيانات مدار البحث ، إن هذا المفهوم يمكن تعميمه عند دراسة أكثر من متغيرين ، فإذا افترضنا أن العلاقة ما بين المتغير التابع Y والمتغيرين المستقلين X_1 و X_2 يمكن وصفها بمعادلة الانحدار المتعدد التالية :

$$\mu_{y/\underline{x}} = \beta_0 + \beta_1 x_1 + \beta_2 x_2$$

والتي يمكن تقديرها من بيانات العينة العشوائية (y_i , x_{i2} , x_{i1}) حيث $i = 1,2,3, \ldots , n$ باستخدام طريقة المربعات الصغرى وبالتالي تكون معادلة الانحدار التقديرية كالآتي :

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2$$

إن معامل التحديد وكما أشرنا يعبر عن نسبة إجمالي الاختلاف في قيم المتغير Y التي يمكن تفسيرها بالنموذج الموفق ويتم تعريف معامل التحديد في هذه الحالة كما يلي :

$$R^2 = 1 - \frac{SS\,Res.}{SST} \quad , 0 \le R^2 \le 1 \qquad (43)$$

حيث $SST = \sum_{i=1}^{n} (y_i - \bar{y})^2$ و $SS\,Res. = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$

ولكن صيغة .SSRes غالباً ما تكون صعبة الاستعمال من الناحية التطبيقية وبالتالي يفضل استخدام الصيغة الآتية في حسابها أي أن :

$$SS\,Res. = \sum_{i=1}^{n} y_i^2 - \hat{\beta}_0 \sum_{i=1}^{n} y_i - \hat{\beta}_1 \sum_{i=1}^{n} x_{i1} y_i - \hat{\beta}_2 \sum_{i=1}^{n} x_{i2} y_i$$

مثال (8) : من بيانات مثال (3) أوجد R^2 .

الحل :

حيث أن $SS\,Res. = 3.78$ و $SST = 321.24$ ، وعليه فإن

$$R^2 = 1 - \left(\frac{3.78}{321.24} \right) = 1 - 0.0118 = 0.9882$$

أي أن وزن الطفل وعمره يفسران 98.82 ٪ من إجمالي الاختلاف في طول الطفل .

إن R^2 تُعد إحصاءة تساعد في تحديد مقدار جودة النموذج في توفيق البيانات ، وعليه كلما كانت قيمته كبيرة كلما دل ذلك على أن النموذج يوفق البيانات بشكل جيد والعكس صحيح .

إن معامل الارتباط المتعدد هو الجذر التربيعي لمعامل التحديد المتعدد ، ويعتبر مقياس لدرجة العلاقة ما بين المتغير Y ومجموعة المتغيرات المستقلة $X_1, X_2, X_3, ..., X_k$ التي يتضمنها نموذج الانحدار المتعدد ، وغالباً ما نرغب في معرفة ما هي المتغيرات التي يجب استبعادها من النموذج وذلك بسبب أنه لا تأثير لها أو أن تأثيرها بسيط في التنبؤ بقيم Y ، أو نود معرفة المتغيرات التي يجب إضافتها لنموذج الانحدار حتى يمكن الحصول على تنبؤ جيد لقيم Y .

حيث أن معامل التحديد يعتبر مقياس للعلاقة ما بين المتغير التابع Y والمتغير المستقل (X_i)، وبالتالي يمكن استخدامه لتحديد أهمية وجود متغير مستقل معين في معادلة الانحدار ، وبالتأكيد إن قوة العلاقة ما بين Y والمتغير المستقل (X_i) مثلاً ، قد يرجع ذلك إلى قوة العلاقة ما بين كلاً من Y و (X_i) مع (X_j) وبالتالي فإن الارتباط الحقيقي ما بين Y و (X_i) يمكن معرفته من خلال إزالة تأثير (X_j) ، ويمكن معرفة ذلك من خلال استخدام معامل الارتباط الجزئي (Partial correlation coefficient) ، ويرمز لهذا المعامل بالرمز $r_{Yi.j}$ وهو يقيس العلاقة ما بين Y و X_i عندما تكون X_j ثابتة وهو معرف كما يلي :

$$r_{Yi.j} = \frac{r_{Yi} - r_{Yj} \quad r_{ij}}{\sqrt{(1 - r_{Yj}^2)(1 - r_{ij}^2)}} \qquad (44)$$

حيث r_{Yi} تمثل معامل الارتباط ما بين Y و X_i

و r_{Yj} تمثل معامل الارتباط ما بين Y و X_j

و r_{ij} تمثل معامل الارتباط ما بين X_i و X_j .

وبالمثل يمكن تعريف $r_{Yj.i}$ الذي يقيس العلاقة ما بين Y و X_j عندما تكون X_i ثابتة . ألحظ أن $r_{Yi.j}^2$ يرمز لمعامل التحديد الجزئي (Coefficient of partial determination) وهو يمثل التناسب ما بين الاختلاف غير المفسر والاختلاف غير المفسر السابق ، وبعبارة أخرى $r_{Yi.j}^2$ تعطي نسبة الاختلاف غير المفسر في قيم Y بمعادلة الانحدار التي تتضمن X_j فقط والذي يمكن تفسيره بعد إضافة X_i مع X_j لنموذج الانحدار .

مثال (9): بفرض أن البيانات التالية تم الحصول عليها من تجربة تهدف لمعرفة إمكانية التنبؤ بوزن (y_i) نوع معين من الحيوانات بعد فترة زمنية معينة وذلك على أساس معرفة وزنه الابتدائي (X_1) وكمية الغذاء (X_2) التي تناولها في هذه الفترة :

$$\sum_{i=1}^{10} y_i = 825 \quad , \quad \sum_{i=1}^{10} x_{i1} = 379 \quad , \quad \sum_{i=1}^{10} x_{i2} = 2417$$

$$\sum_{i=1}^{10} y_i^2 = 70083 \quad , \quad \sum_{i=1}^{10} x_{i1}^2 = 14533 \quad , \quad \sum_{i=1}^{10} x_{i2}^2 = 601365 \quad ,$$

$$\sum_{i=1}^{10} x_{i1} y_i = 31726 \quad , \quad \sum_{i=1}^{10} x_{i2} y_i = 204569 \quad , \quad \sum_{i=1}^{10} x_{i1} x_{i2} = 92628$$

والمطلوب إيجاد كلاً من :

ا - معامل الارتباط المتعدد .

ب - ما هي نسبة التخفيض في الاختلاف في وزن الحيوان عند إضافة كمية الغذاء لمعادلة الانحدار $(r_{Y2.1})$.

الحل :

ا - لإيجاد معامل الارتباط المتعدد يجب أولاً إيجاد قيم معلمات الانحدار التقديرية أي الحصول على معادلة الانحدار التقديرية

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2$$

ومن المعطيات أعلاه فإن المعادلات الطبيعية ستكون كالآتي :

$$10\hat{\beta}_0 + 379\hat{\beta}_1 + 2417\hat{\beta}_2 = 825$$

$$379\hat{\beta}_0 + 14533\hat{\beta}_1 + 92628\hat{\beta}_2 = 31726$$

$$2417\hat{\beta}_0 + 92628\hat{\beta}_1 + 601365\hat{\beta}_2 = 204569$$

وبحل هذه المعادلة نجد أن

$\hat{\beta}_2 = 0.218$ و $\hat{\beta}_1 = 1.396$ و $\hat{\beta}_0 = -22.992$

وعليه فإن معادلة الانحدار التقديرية تكون كالآتي :

$$\hat{y} = -22.992 + 1.396\, x_1 + 0.218\, x_2$$

وإن

$$SS\,Res. = \sum_{i=1}^{10} y_i^2 - \hat{\beta}_0 \sum_{i=1}^{10} y_i - \hat{\beta}_1 \sum_{i=1}^{10} x_{i1} y_i - \hat{\beta}_2 \sum_{i=1}^{10} x_{i2} y_i$$

$$= 70083 - (-22.992)(825) - (1.396)(31726) - (0.218)(204569)$$

$$= 165.862$$

$$SST = \sum_{i=1}^{10}(y_i - \bar{y})^2 = \sum_{i=1}^{10} y_i^2 - \frac{\left(\sum_{i=1}^{10} y_i\right)^2}{10}$$

$$= 70083 - \frac{(825)^2}{10} = 2020.5$$

بذن

$$R^2 = 1 - \frac{SSRes}{SST} = 1 - \left(\frac{165.862}{2020.5}\right)$$

$$= 1 - 0.082 = 0.918$$

ب -

$$r_{Y1} = \frac{n\sum_{i=1}^{n} x_{i1} y_i - \left(\sum_{i=1}^{n} x_{i1}\right)\left(\sum_{i=1}^{n} y_i\right)}{\left\{\left[n\sum_{i=1}^{n} x_{i1}^2 - \left(\sum_{i=1}^{n} x_{i1}\right)^2\right]\left[n\sum_{i=1}^{n} y_i^2 - \left(\sum_{i=1}^{n} y_i\right)^2\right]\right\}^{\frac{1}{2}}}$$

$$r_{Y1} = \frac{(10)(31726) - (379)(825)}{\left\{\left[(10)(14533) - (379)^2\right]\left[(10)(70083) - (825)^2\right]\right\}^{\frac{1}{2}}}$$

$$= \frac{4585}{\left[(1689)(20213)\right]^{\frac{1}{2}}}$$

$$= \frac{4585}{5841.767} = 0.785$$

وان

$$r_{Y2} = \frac{n\sum_{i=1}^{n} x_{i2} y_i - \left(\sum_{i=1}^{n} x_{i2}\right)\left(\sum_{i=1}^{n} y_i\right)}{\left\{\left[n\sum_{i=1}^{n} x_{i2}^2 - \left(\sum_{i=1}^{n} x_{i2}\right)^2\right]\left[n\sum_{i=1}^{n} y_i^2 - \left(\sum_{i=1}^{n} y_i\right)^2\right]\right\}^{\frac{1}{2}}}$$

$$r_{Y2} = \frac{(10)(204569) - (2417)(825)}{\left\{\left[(10)(601365) - (2417)^2\right]\left[(10)(70083) - (825)^2\right]\right\}^{\frac{1}{2}}} = 0.877$$

وان

$$r_{12} = \frac{n\sum_{i=1}^{n} x_{i1}x_{i2} - \left(\sum_{i=1}^{n} x_{i1}\right)\left(\sum_{i=1}^{n} x_{2i}\right)}{\left\{\left[n\sum_{i=1}^{n} x_{i1}^2 - \left(\sum_{i=1}^{n} x_{i1}\right)^2\right]\left[n\sum_{i=1}^{n} x_{i2}^2 - \left(\sum_{i=1}^{n} x_{i2}\right)^2\right]\right\}^{\frac{1}{2}}}$$

$$r_{12} = \frac{(10)(92628)-(379)(2417)}{\left\{\left[(10)(14533)-(379)^2\right]\left[(10)(601365)-(2417)^2\right]\right\}^{\frac{1}{2}}}$$

$$= \frac{50237}{\left[(1689)(171761)\right]^{\frac{1}{2}}} = \frac{50237}{17032.449} = 0.601$$

وعليه فإن

$$r_{Y2.1} = \frac{r_{Y2} - r_{Y1}\, r_{12}}{\sqrt{(1-r_{Y1}^2)\,(1-r_{12}^2)}}$$

$$= \frac{0.877-(0.785)(0.601)}{\sqrt{\left[1-(0.785)^2\right]\left[1-(0.601)^2\right]}}$$

$$= \frac{0.877-0.472}{\sqrt{(0.384)(0.639)}} = \frac{0.405}{0.495} = 0.818$$

حيث أن $r_{Y2.1}^2 = 0.669$ وعليه فإن إضافة كمية الغذاء التي يتناولها الحيوان للنموذج قد أحدث تخفيض قدره 66.9 ٪ في الاختلاف في وزن الحيوان غير المفسر بمعادلة الانحدار باستخدام الوزن الابتدائي فقط .

تمرينات Exercises

1- من البيانات التالية :

x	4	5	9	14	18	22	24
y	16	22	11	16	7	3	17

أوجد كلا من :

أ - معامل الارتباط وفسر الناتج .

ب - عند مستوى المعنوية 5 ٪ أختبر الفرضية $H_o:\rho = 0$ مقابل جميع البدائل الممكنة .

جـ - عند مستوى المعنوية 1 ٪ أختبر الفرضية $H_o:\rho = -0.8$ مقابل جميع البدائل الممكنة .

2- الجدول التالى يمثل اوزان كل من القلب (x) و الكبد (y) لعشر فئران :

x	1.3	1.5	1.7	1.4	2.0	1.3	1.5	1.8	1.6	1.9
y	18	17	24	16	28	17	19	27	20	24

أ - ارسم الشكل الانتشارى .
ب - اوجد معامل الارتباط مع التعليق على الناتج .

جـ - اوجد معادلة انحدار y على x .
د - احسب تباين الانحدار .

هـ - اوجد قيمة y التقديرية عندما x = 2.2 .

3 - البيانات الاتية توضح المتوسط الحسابى و الانحراف المعيارى لعمر (x) و ضغط الدم (y) لعشر اصحاء فاوجد تقديراً لضغط دم شخص ما كان عمره 50 سنة .

	العمر (x)	ضغط الدم (y)
المتوسط الحسابى :	52	145
الانحراف المعيارى :	12.5	14

علماً بأن :

$$\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) = 1240$$

4 - للبيانات التالية :

x : 1	2	3	4	5	6
y : 6	4	3	5	4	2

أ - ارسم الشكل الانتشارى .

ب - اوجد معامل الارتباط بين المتغيرين x و y .

جـ - اوجد معادلة انحدار y على x التقدرية .

د - اوجد تقدير للمعلمة $\mu_{y \setminus x=4}$

هـ - اختبر الفرضية التالية عند مستوى المعنوية 0.05 .

$H_0 : \beta_1 = 0$ vs $H_1 : \beta_1 \neq 0$.

و - اختبر الفرضية التالية عند مستوى المعنوية 0.01 .

$H_0 : \beta_0 = 0$ vs $H_1 : \beta_0 \neq 0$.

5 - إذا علمت ان :

$$\sum_{i=1}^{25} x_i^2 = 625 \quad , \quad \sum_{i=1}^{25} y_i = 100 \quad , \quad \sum_{i=1}^{25} x_i = 125$$

$$\sum_{i=1}^{25} x_i y_i = 508 \quad , \quad \sum_{i=1}^{25} y_i^2 = 460$$

أ - اوجد قيمة معامل الارتباط بين المتغيرين x و y .

ب - اوجد معادلة انحدار y على x التقدرية .

جـ - إذا اكتشف بأن القيمتين (6 ، 14)، (8 ، 6) حسبتا عن طريق الخطأ بدلاً من (8 ،12) (6 ، 14) فما قيمة معامل الارتباط الصحيح .

د - أوجد معادله أنحدار y على x التقديرية بعد عملية التعديل .

هـ - أوجد قيمة المتغير التابع التقديرية عندما x=15 .

6 - البيانات التالية تمثل درجات 9 طلبه في الامتحان النصفى (x) والامتحانات النهائى (y) في أحد المقررات الدراسية .

x : 67	99	96	94	81	72	71	50	77
y : 68	99	99	85	47	34	78	66	82

أ. أوجد معادلة الانحدار التقديرية .

ب. أوجد تقديراً لدرجة الامتحان النهائى لطالب درجته في الامتحـان النصفـى 85 و لكنـه تغيب عن الامتحان النهائى بسبب حاله مرضية .

7 - إذا علمت بأن معادلتى الانحدار للمتغيرين x و y هى

$40x-18y=214$ ، $8x-10y+66=0$

فإذا كان تباين المتغير X يساوى 9 و فاوجد :

أ. متوسطى x و y .   ب - معامل الارتباط بين المتغيرين x و y .   جـ - تباين المتغير y .

8 - البيانات التالية تمثل العمر الزمنى بالسنوات (x) لنـوع معيـن مـن السـيارات و أسـعارها بالدينار (y)

العمر الزمنى :	1	2	2	3	5	5
القيمة بالدينار :	2350	1695	1750	1395	985	895

أ. وفق هذه البيانات بإستخدام المنحنى الاسى . (ارشاد : $\mu_{y\backslash x}=\beta_0\beta^x$) .

ب. أوجد تقديراً لسعر بيع سياره عمرها الزمنى 4 سنوات .

9 - للبيانات التالية :

x :	1	2	3	4	5	6	7	8	9
y :	9.1	7.3	3.2	4.6	4.8	2.9	5.7	7.1	8.8

أ - وفق منحنى الانحدار الذى معادلته : $\mu_{y\backslash x}=\beta_0+\beta_1x+\beta_2x^2$.

ب - اوجد تقدير لقيمة y عندما $x=2$.

10 - حدد فما إذا كانت النماذج التالية خطيه في المعلمات أو المتغيرات أو كلاهما :

أ - $y_i = \beta_0 + \beta_1 \frac{1}{x_i} + \varepsilon_i$.

ب - $\mathrm{Log}_e Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$.

جـ - $Y_i = \beta_0 + \beta_1 \mathrm{Log}_e X_i + \varepsilon_i$.

د - $\mathrm{Log}_e Y_i = \mathrm{Log}_e \beta_0 + \beta_1 \mathrm{Log}_e X_i + \varepsilon_i$.

هـ - $Y_i = \beta_0 + (0.75 - \beta_0) e^{-\beta_1 (x_i - 2)} + \varepsilon_i$.

و - $Y_i = \beta_0 + \beta_1^3 x_i + \varepsilon_i$.

11 - أثبت أن : $0 \le r_{xy}^2 \le 1$

(ارشاد : استخدم متباينه كوشى - شوارز .

$$\left(\left[E(xy)\right]^2 \le E\left(x^2\right)E\left(y^2\right)\right)$$

12 - أثبت أن : $r^2 = \hat{\beta}_{yx} \hat{\beta}_{xy}$ حيث :

$\hat{\beta}_{xy}$ = ميل معادلة انحدار y على x التقديرية .

$\hat{\beta}_{yx}$ = ميل معادلة انحدار x على y التقديرية .

r = معامل الارتباط بين المتغيرين x و y .

13 - إذا كان :

$$X_i^* = \frac{x_i - \bar{x}}{S_x} \qquad , \qquad Y_i^* = \frac{y_i - \bar{y}}{S_y}$$

حيث S_x و S_y يمثلان الانحراف المعيارى للمتغيرين X و Y على الترتيب .فالنمودج التالى :

$$X_i^* = \alpha + \beta y_i^* + \varepsilon_i$$

أثبت أن : $\hat{\beta} = r_{xy}$, $\hat{\alpha} = 0$.

14 - في معادلة الانحدار الخطى البسيط إذا كان خط الانحدار يمر بنقطة الاصل فإن النمودج يكون علي الصورة التالية : $Y_i = \beta x_i + \varepsilon_i$.

و المطلوب أثبات أن :

أ- $\hat{\beta} = \dfrac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2}$.

ب- $Var(\hat{\beta}) = \dfrac{\sigma^2}{\sum_{i=1}^{n} x_i^2}$.

جـ- $\sum_{i=1}^{n} e_i$. ليس من الضرورى أن تساوى صفر .

15 - إذا كان :

$$(I)\quad Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i .$$

$$(II)\quad Y_i = \alpha_0 + \alpha_1 \left(X_i - \overline{X}\right) + \varepsilon_i .$$

أ- أوجد مقدرى α_0 و β_0 ؟ هل هما متطابقان ؟ و هل لهما نفس التباين ؟

ب- أوجد مقدرى α_1 و β_1 ؟ هل هما متطابقان ؟ و هل لهما نفس التباين ؟

جـ- ما هى ميزة (إن وجدت) النمودج الثانى عن الاول ؟

16 - إذا كان :

$$(I)\quad Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$$

$$(II)\quad Y_i^* = \alpha_0 + \alpha_1 X_i^* + \varepsilon_i .$$

حيث : $Y_i^* = \dfrac{Y_i - \overline{Y}}{S_y}$ ، $X_i^* = \dfrac{X_i - \overline{X}}{S_x}$ أثبت أن : $\hat{\alpha}_1 = \hat{\beta}_1 \left(\dfrac{S_x}{S_y}\right)$.

17 - إذا كانت r_1 تمثل معامل الارتباط بين n من أزواج القيم (X_i, Y_i) و r_2 تمثل معامل الارتباط بين n من أزواج القيم $(aX_i + b, cY_i + d)$ حيث a, b, c, d ثوابت . أثبت أن $r_1 = r_2$.

18 - إذا كان معامل الارتباط بين n من أزواج القيم (X_i , Y_i) موجباً. فبرهن صحة ما يلى :

أ - الارتباط بين $(-X_i , -Y_i)$ يكون موجباً .

ب - الارتباط بين $(-X_i , Y_i)$ و كذلك الارتباط بين (X_i , Y_i) قد يكون موجباً أو سالباً .

جـ - معامل الانحدار β_{xy} و β_{yx} كلاهما موجب .

19 - للبيانات التالية :

x	88	70	65	50	60	80	68	49	40
y	95	51	49	27	42	52	67	48	46

أ - ارسم الشكل الانتشارى .

ب - أوجد معادلة انحدار y على x التقديرية .

جـ - احسب الانحراف المعيارى للقيم التقديرية لمعاملات الانحدار .

د - كون فترة بدرجة ثقة 95٪ حول كلا من β_0 ، β_1 و σ^2 .

هـ - اختبر الفرضيات التالية عند مستوى معنويه 0.05 .

1 - $H_0 : \beta_0 = 0 \quad vs \quad H_1 : \beta_0 \neq 0$.

2 - $H_0 : \beta_1 = 0 \quad vs \quad H_1 : \beta_1 \neq 0$.

و - اختبر الفرضية (2) فى الفقرة (هـ) بإستخدام جدول تحليل البيانات ثم قارن بين النتيجتين

20 - للبيانات التالية :

x :	-5	-4	-3	-2	-1	0	1	2	3	4	5
y :	1	5	4	7	10	8	9	13	14	13	18

بإفتراض أن النموذج المناسب لتوفيقها هو :

$$Y = \beta_0 + \beta_1 x + \varepsilon \ .$$

أ - ما هى تقديرات المربعات الصغرى لكل من β_0 و β ؟ ثم أوجد المعادلة التقديرية .

ب - كون جدول تحليل التباين ثم اختبر الفرضية $H_0 : \beta_1 = 0$ عند مستوى معنوية 0.05 .

جـ - كون فترة بدرجة ثقة 95٪ حول β_1 .

د - كون فترة بدرجة ثقة 95٪ حول المتوسط الفعلى للمتغير y عندما $x = 3$.

هـ - كون فترة بدرجة ثقة 95٪ حول الفرق بين المتوسط الفعلى للمتغير y عندما $x = 3$ و المتوسط الحقيقى للمتغير y عندما $x = -2$.

21 - للبيانات الآتية :

x_1	x_2	y
1	8	6
4	2	8
9	− 8	1
11	− 10	0
3	6	5
8	− 6	3
5	0	2
10	− 12	− 4
2	4	10
7	− 2	− 3
6	− 4	5

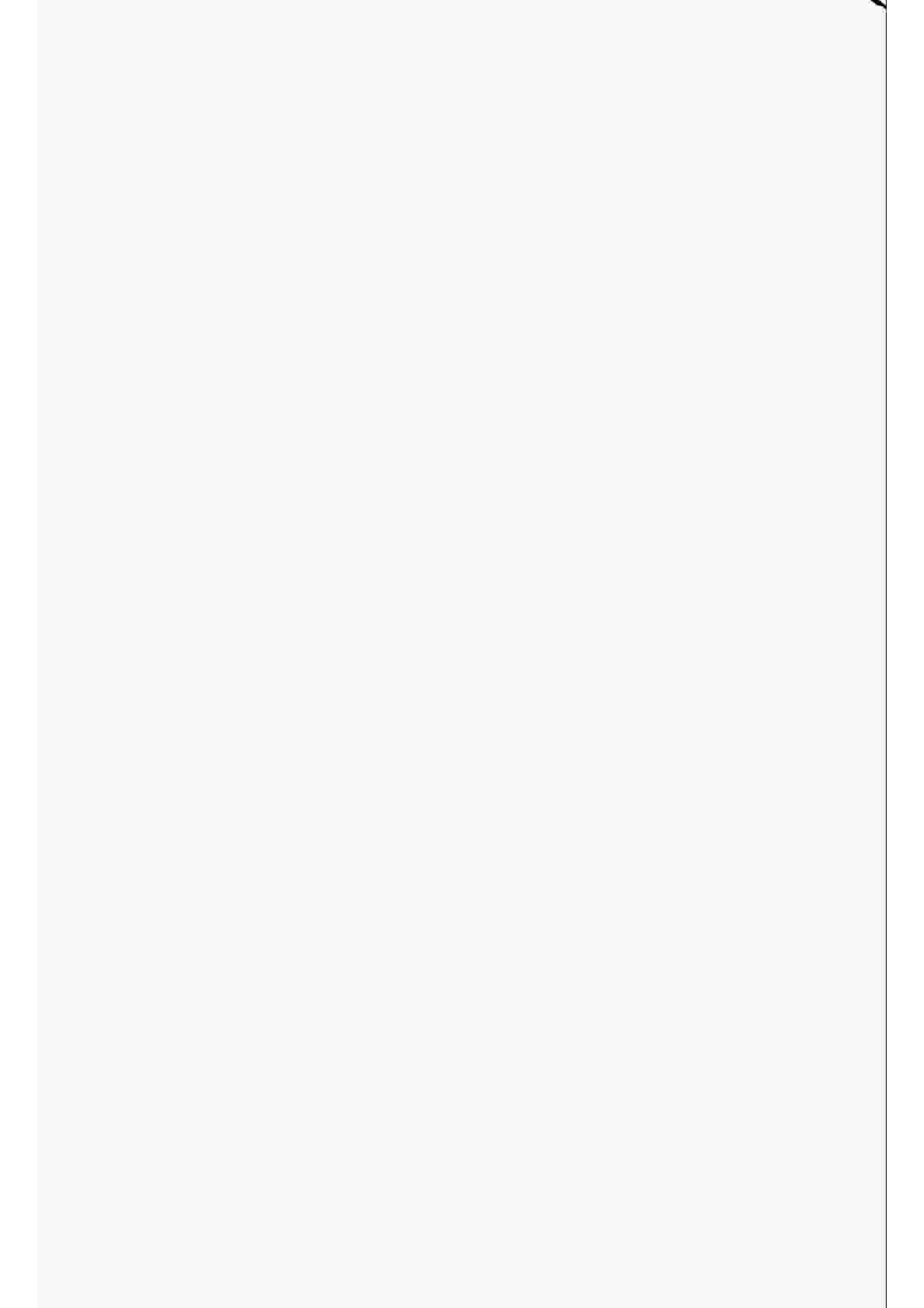
أ - بإستخدام طريقة المربعات الصغرى أوجد القيم التقديرية لمعاملات الانحدار فى النموذج التالى:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon \ .$$

ب - كون جدول تحليل التباين .

جـ - اختبر معنوية الانحدار مستخدماً $\alpha = 0.05$.

د - أوجد قيمة معامل التحديد مع التعليق على الناتج .



الفصل العاشر

تحليل التبـــاين

Analysis of Variance

10-1 مقدمـــــة Introduction

يعتبر اختبار التوزيع الطبيعي من أقوى الاختبارات الإحصائية للمقارنة بين متوسطي مجتمعين عندما يكون تباين المجتمعين معلوماً بينما يعتبر اختبار t هو الأقوى عندما يكون تباين المجتمعين مجهولاً ولكنهما متساويين والمعاينة من مجتمعات طبيعية ، ولكن إذا زاد عدد المقارنات عن أثنين فإن اختبار t يصبح أداة غير عملية لأننا سنجرى اختبار تساوى متوسطات هذه المجتمعات مثنى مثنى ، بالإضافة إلى ذلك إن احتمال رفض فرضية صحيحة يكون أعلى بكثير من مستوى المعنوية المعلن عنه . فمثلاً إذا كان لدينا 4 عينات سيكون هناك 6 اختبارات وإذا كانت $\alpha=0.05$ فإن احتمال أتخاذ قرار صحيح هو 0.95 للاختبار الواحد ويكون تبعاً لذلك احتمال القرار الصحيح للاختبارات الستة هو $(0.95)^6$ ،وبذلك يكون احتمال اتخاذ قرار خاطئ على الأقل في اختبار واحد من هذه الاختبارات هو $1-(0.95)^6=0.26$ ، أي أن 0.26 من الحالات سوف نقع في ارتكاب النوع الأول من الخطأ . عليه فقد وجدت طريقة أخرى تستخدم في مثل هذه الحالات هذه الطريقة سميت بتحليل التباين (Analysis of Variance) والذي يمكن تعريفه كالأتي :

هو أسلوباً إحصائياً يمكن بواسطته تجزئة أجمالي التباين الموجود في مجموعة من البيانات إلى عدة عناصر ،مصاحب لكل منها مصدر معين من التباين ،وبواسطته يمكن تحديد مقدار مشاركة كل مصدر من هذه المصادر في أجمالي التباين .

ولتوضيح بعض المفاهيم المستخدمة في تحليل التباين سوف نبدأ بدراسة المثال الأتي : لنفرض أن باحث بمركز البحوث الحيوانية يرغب في مقارنة ثلاث أنواع من الفيتاميات A و B و C وذلك لمعرفة مدى تأثير كل منها في زيادة وزن نوع معين من الحيوانات خلال فترة زمنية معينة من تناولها، وللوصول إلى هذا الهدف قام باختيار عينة عشوائية من هذه الحيوانات وأعطى مجموعة منها الفيتامين A، ولمجموعة أخرى الفيتامين B، ولمجموعة ثالثة الفيتامين C، وبعد فترة زمنية معينة قام بوزن هذه الحيوانات فوجد أن هناك اختلاف في مقدار الزيادة في الوزن بين المجموعات الثلاث ، وأيضاً هناك اختلاف في مقدار الزيادة في الوزن داخل نفس المجموعة التي أعطى لها نفس النوع من الفيتامين ، أي أنه هناك نوعين من المتغيرات هما : متغير المعالجة

مختلفة (مستويات) وله ثلاث قيم (العامل) والذي يمثله الفيتامين (treatment variable) وهي A و B و C ، أما المتغير الآخر فهو المتغير التابع (أو المستجيب Response) variable والذي يمثله مقدار الزيادة في الوزن أي أن المتغير المستجيب هو المتغير الذي نتوقع أن يعطى قيم مختلفة عند تطبيق قيم مختلفة لمتغير المعالجة عليه ، وما يهم الباحث هنا هو متغير المعالجة ، والسؤال الذي يرغب في الإجابة عليه هو : هل أن القيم المختلفة لمتغير المعالجة سيؤدى في المتوسط إلى نتائج مختلفة للمتغير المستجيب ؟ . للإجابة على هذا السؤال يتطلب الأمر تحليل الاختلاف الكلى المشاهد في الوزن إلى مكوناته المختلفة ويتم ذلك من خلال استخدام أسلوب تحليل التباين . إن أسلوب تحليل التباين له استخدامات عديدة في كثير من المجالات ، فمثلاً عندما نود معرفة فيما إذا كان هناك فروق معنوية بين عدة أنواع مختلفة من السماد أو عدة طرائق مختلفة للتدريس أو عدة أنواع من الأدوية أو دراسة تأثير أنواع مختلفة من السماد على أنواع مختلفة من القمح... وهكذا . أي أنه يمكن القول بأن هذا الأسلوب يستخدم لغرضين هما :

1- تقدير واختبارات الفروض الخاصة بمتوسطات المجتمعات الإحصائية .

2- تقدير واختبارات الفروض الخاصة بتباينات المجتمعات الإحصائية .

وإذا كان تحليل التباين يختص بدراسة عامل (متغير) واحد فإنه يسمى بتحليل التباين الأحادي أما إذا كان يهتم بدراسة عاملين فإنه يسمى بتحليل التباين الثنائي ...وهكذا .وسوف نتعرض لأنواع تحليل التباين المختلفة في البنود القادمة .

10 - 2 التصميم العشوائي الكامل The Completely Randomized Design (C.R.D)

إن أبسط أنواع تحليل التباين هو تحليل التباين الأحادي (One- Way Analysis of Variance) الذي يتم من خلاله دراسة مصدر (أو عامل (factor) كما يطلق عليه في بعض الأحيان) واحد من الاختلاف ، إن التجربة التي يتم تحليلها باستخدام أسلوب تحليل التباين الأحادي يتم تصميمها بحيث أن المعالجات (مستويات العامل) تصنف بطريقة عشوائية كاملة للوحدات التجريبية (experimental units) التي سيتم منها أخذ القياسات وذلك لغرض معرفة تأثير المعالجات عليها، ولهذا السبب يطلق على مثل هذا النوع من التصميم تسمية التصميم العشوائي الكامل (completely randomized design) ، مع مراعاة أن تكون الوحدات التجريبية متجانسة ، ولتوضيح فكرة هذا التصميم لنفرض أنه لدينا 12 حيوان مشارك في تجربة وذلك لغرض المقارنة بين ثلاثة أنواع من الفيتامينات وهى A و B و C تستخدم في تغذية هذه الحيوانات . إن أول خطوة في هذا التصميم هي ترقيم هذه الحيوانات من 01 إلى 12 ثم استخدام جدول الأرقام

العشوائية واختيار 12 رقم متتالي ، وحيث أن أكبر رقم أعطي لترقيم الحيوانات هو 12 ومكون من خانتين ، وعليه نختار عمودين من جداول الأرقام العشوائية ونختار 12 رقم محصور ما بين 1 و 12 ، وذلك مع إهمال كل عدد اكبر من 12 أو اي رقم متكرر علماً بأن نقطة البداية بهذا الجدول اختيارية ، وبالرجوع إلى الجدول الذي بأخر هذا الكتاب وكانت نقطة البداية بالصف 10 والعموديين 9 و 10 فإن الرقم عند هذا التقاطع هو 39 وعليه وبالتحرك إلى أسفل فإن الأرقام التي سيتم اختيارها هي : 05 ، 10 ، 04 ، 12 ، 03 ، 06 ، 02 ، 09 ، 11 ، 07 ، 01 ، 08 . وهذا يعنى أن الحيوانات التي أرقامها 05 ، 10 ، 04 ، 12 سيتم تغذيتها بالفيتامين A ، والحيوانات التي أرقامها 03 ، 06 ، 02 ، 09 سيتم تغذيتها بالفيتامين B ، والحيوانات التي أرقامها 11 ، 07 ، 01 ، 08 ، سيتم تغذيتها بالفيتامين C ، وبالتالي يكون شكل التصميم النهائي كما يلي :

C 1	B 2	B 3
A 4	A 5	B 6
C 7	C 8	B 9
A 10	C 11	A 12

لحظ أنه إذا كان عدد الوحدات التجريبية غير متساوى بكل معالجة فإننا نتبع نفس الأسلوب مع مراعاة عدد الوحدات بكل معالجة فقط .

إذن إذا كان التصميم العشوائى الكامل هو التصميم المناسب للتجربة وكان هناك t من المعالجات (أو المجتمعات) مثل t نوع من الأدوية ، أو من الأسمدة الزراعية ، أومن طرق التدريس المختلفة ...الخ. ، وكان بكل معالجة n من المشاهدات (أو القراءات أو القياسات أو التكرارت replicates) فإن نتائج هذا التصميم تكون كما هي موضحة في الجدول التالي :

	المعالجات						
	1	2	...	i	...	t
	y_{11}	y_{21}	...	y_{i1}	...	y_{t1}
	y_{12}	y_{22}	...	y_{i2}	...	y_{t2}
	⋮	⋮		⋮		⋮
	y_{1n}	y_{2n}	...	y_{in}	...	y_{tn}
المجموع	$y_{1.}$	$y_{2.}$	...	$y_{i.}$	...	$y_{t.}$
المتوسط	$\bar{y}_{1.}$	$\bar{y}_{2.}$	...	$\bar{y}_{i.}$	...	$\bar{y}_{t.}$

حيث y_{ij} ترمز للمفردة j الناتجة من المعالجة i و $i = 1,2,\cdots,t$ و $j = 1,2,\cdots,n$. وذلك على افتراض أن عدد المفردات (المشاهدات) في المعالجات متساوى ، وإن $y_{i.} = \sum_{j=1}^{n} y_{ij}$ ترمز لمجموع المفردات التى بالمعالجة i ، و $\bar{y}_{i.} = \frac{y_{i.}}{n}$ ترمزلمتوسط المفردات التى بالمعالجة i ، و $y_{..} = \sum_{i=1}^{t} y_{i.} = \frac{\sum_{i=1}^{t}\sum_{j=1}^{n} y_{ij}}{N}$ ترمز لمجموع جميع المفردات وهى n t مفردة ، و $\bar{y}_{..} = \frac{y_{..}}{N}$ لمتوسط جميع المفردات ($N = nt$) .

حيث أن كل مفردة داخل كل معالجة (أو مجتمع) يمكن أن تساوى المتوسط الحقيقي للمعالجة (μ_i) مضافاً إليه مقدار أخر هذا المقدار قد يكون صفراً أو كمية موجبة أو سالبة ، وهذا يعنى أنه من الممكن وجود اختلاف ما بين المفردة بمعالجة ما ومتوسط مفردات تلك المعالجة، وبالتالي يطلق على هذا الاختلاف تسمية الخطأ العشوائي ويرمز له بالرمز ϵ_{ij} . إن المقصود بكلمة الخطأ هنا ليس المعنى المألوف لها وإنما نقصد بذلك الاختلافات الخارجية الموجودة ما بين عناصر أو مفردات أى مجتمع ، وعليه إذا تمت إضافة ϵ_{ij} لمتوسط أى مجموعة (μ_i) فإن الناتج سيكون المفردة (y_{ij}) المختلفة عن متوسطها بمقدار ϵ_{ij} ، وبالتالي يمكن كتابة كل مفردة على الصيغة الآتية :

$$y_{ij} = \mu_i + \epsilon_{ij} \quad (1)$$

وعليه من المعادلة (1) نجد أن : $\in_{ij} = y_{ij} - \mu_i$ ، وحيث أنه هناك t من المجتمعات (المعالجات)

فإن المتوسط العام لجميع المفردات في جميع المجتمعات (المعالجات) نرمز له بالرمز μ حيث

$$\mu = \frac{\sum_{i=1}^{t} \mu_i}{t} .$$

وكما في حالة اختلاف أى مفردة داخل أى معالجة (مجتمع) عن متوسط مفردات تلك المعالجة بمقدار معين ، فإنه من الممكن أن يختلف متوسط أى معالجة (مجتمع) عن المتوسط العام لجميع المعالجات (المجتمعات) بمقدار معين هذا المقدار من الاختلاف يطلق عليه تسمية تأثير المعالجة (treatment effect) ويرمز لتأثير المعالجة i بالرمز α_i أى أن : $\alpha_i = \mu_i - \mu$

وبالتالي فإن :

$$\mu_i = \mu + \alpha_i \tag{2}$$

وبتعويض (2) في (1) نجد أن :

$$y_{ij} = \mu + \alpha_i + \in_{ij} \tag{3}$$

إن النموذج (3) يطلق عليه تسمية نموذج تحليل التباين الأحادي وذلك لأن المعالجات (أو المجتمعات والتى عددها t) تحتوى اختلاف في عامل واحد فقط مثل : أدوية مختلفة ، أسمدة مختلفة ، طرائق تدريس مختلفة ، ... الخ. ، ويقوم هذا النموذج على الافتراضات التالية :

1 - إن جميع المشاهدات تشكل t من العينات العشوائية المستقلة من مجتمعاتها المختارة منها .

2 - كل مجتمع من هذه المجتمعات يتوزع توزيعاً طبيعياً بمتوسط يساوى μ_i وتباين يساوى σ_i^2 وإن جميع التباينات متساوية أى أن $\sigma_1^2 = \sigma_2^2 = ... = \sigma_t^2 = \sigma^2$.

3 - إن تأثير جميع المعالجات α_i و $i = 1,2,\cdots,t$ مقادير ثابتة ، وحيث أن $\alpha_i = \mu_i - \mu$

وعليه فإن $\sum_{i=1}^{t} \alpha_i = 0$.

4 - حيث أن $\in_{ij} = y_{ij} - \mu_i$ وعليه فإن $\in_{ij}$ متغيرات عشوائية ومستقلة ولكل منها توزيع طبيعي ، أى أن $\in_{ij} \sim NID(0, \sigma^2)$.

لقد أشرنا في البداية إلى أنه من أهداف تحليل التباين هو التقدير واختبار الفرضيات الإحصائية ، وعليه بدلالة المعلمات التى قدمناها سابقاً يمكن كتابة متوسط الوحدة التجريبية j التى أعطى لها المعالجة i كما يلي :

$$E(Y_{ij}) = \mu + \alpha_i \quad ,i = 1,2,\cdots,t \tag{4}$$

ويمكن أيجاد مقدرات لمعلمات نموذج تحليل التباين الأحادي (μ و α_i)، وإن المعيار المناسب للحصول على تقديرات جيدة هو أن يكون مجموع مربعات الأخطاء أو الانحرافات $\in_{ij}$ أصغر ما يمكن ، هذه الطريقة تسمى بطريقة المربعات الصغرى (l.s.m) التي سبق وأن أشرنا إليها في فصل سابق ، وعند استخدام هذه الطريقة ليس بالضرورة أن يكون للأخطاء توزيع طبيعي، وبما أننا عرفنا تأثير المعالجات كانحرافات عن المتوسط العام μ وعليه فإنه بتطبيق القيد $\sum_{i=1}^{t}\alpha_i = 0$ فإن التقدير بقيمة واحدة لهذه المعلمات وفقاً لطريقة المربعات الصغرى يكون كما يلي :

$$\hat{\mu} = \bar{y}_{..}$$

$$\hat{\alpha}_i = \bar{y}_{i.} - \bar{y}_{..} \quad , \quad i = 1,2,\ldots,t \qquad (5)$$

إن هذا الحل ليس وحيد (unique) وذلك لأنه يعتمد على القيد الذي تم تطبيقه ، ولكن هناك دوال معينة في معلمة النموذج لها تقدير وحيد بصرف النظر عن القيد المطبق على المعادلات الطبيعية ، فمثلاً ، المقدار $\alpha_i - \alpha_j$ يمكن تقديره باستخدام $\hat{\alpha}_i - \hat{\alpha}_j = \bar{y}_{i.} - \bar{y}_{j.}$ والمقدار $\mu + \alpha_i$ يمكن تقديره باستخدام $\hat{\mu} + \hat{\alpha}_i = \bar{y}_{i.}$. وحيث أننا غالباً ما يهمنا هو الفروق ما بين تأثير المعالجات عوضاً عن القيم الفعلية لها ، وعليه فإن عدم وجود حل وحيد لتأثير المعالجة α_i ليس له أهمية كبيرة ، وبصفة عامة إن أى دالة في معلمات النموذج وتشكل تركيبة خطية في α_i و μ يكون لها تقدير وحيد ، وإن الدوال التى لها تقدير وحيد تسمى دوال قابلة للتقدير (estimable function) . إن الهدف الثاني من تحليل التباين وكما أشرنا سلفاً هو اختبار الفرضيات الإحصائية وإن الفرضية الممكن اختبارها هنا هي :

$$H_0 : \mu_1 = \mu_2 = \ldots = \mu_t$$

H_1 : على الأقل متوسطين غير متساويين

أى أن فرض العدم ينص على أن لجميع المعالجات (أو المجتمعات) متوسطات متساوية ، مقابل الفرض البديل الذي ينص على أنه على الأقل متوسطي معالجتين يكونا مختلفين ، وإذا كان لجميع المعالجات نفس المتوسط فهذا يعنى أن تأثير كل معالجة يساوى صفر ، وبالتالي فإن الفرضية السابقة يمكن إعادة كتابتها كالآتي :

$$H_0 : \alpha_1 = \alpha_2 = \alpha_3 = \cdots = \alpha_t = 0$$

H_1 : على الأقل واحدة من α_i لا تساوى صفر

إن الاختبار الذي سوف يستخدم لصحة الفرضية أعلاه سيكون مبنى على مقارنة مقدرين مستقلين لتباين المجتمع المشترك (σ^2) ،أحدهما مبنى على أساس الاختلاف بين متوسطات

لمعالجات والمتوسط العام (بين المعالجات) ، والأخر على أساس الاختلاف ما بين المفردات داخل المعالجات ومتوسط المعالجة (خلال المعالجات) ، ويمكن الحصول على هذين المقدرين من خلال تجزئة الاختلاف الكلي في البيانات وذلك كما يلي :

$$\sum_{i=1}^{t}\sum_{j=1}^{n}(y_{ij}-\bar{y}_{..})^2=\sum_{i=1}^{t}\sum_{j=1}^{n}[(y_{ij}-\bar{y}_{i.})+(\bar{y}_{i.}-\bar{y}_{..})]^2$$

$$=\sum_{i=1}^{t}\sum_{j=1}^{n}[(y_{ij}-\bar{y}_{i.})^2+2(y_{ij}-\bar{y}_{i.})(\bar{y}_{i.}-\bar{y}_{..})+(\bar{y}_{i.}-\bar{y}_{..})^2]$$

$$=\sum_{i=1}^{t}\sum_{j=1}^{n}(y_{ij}-\bar{y}_{i.})^2+2\sum_{i=1}^{t}\sum_{j=1}^{n}(y_{ij}-\bar{y}_{i.})(\bar{y}_{i.}-\bar{y}_{..})$$

$$+\sum_{i=1}^{t}\sum_{j=1}^{n}(\bar{y}_{i.}-\bar{y}_{..})^2$$

وحيث أن

$$\sum_{j=1}^{n}(y_{ij}-\bar{y}_{i.})=y_{i.}-n\bar{y}_{i.}=y_{i.}-n\left(\frac{y_{i.}}{n}\right)=0$$

وأن الحد الثالث لا يحتوى على الرمز j وبالتالي يمكن كتابته كما يلي :

$$\sum_{i=1}^{t}\sum_{j=1}^{n}(\bar{y}_{i.}-\bar{y}_{..})^2=n\sum_{i=1}^{t}(\bar{y}_{i.}-\bar{y}_{..})^2 \qquad (6)$$

وعليه فإن :

$$\sum_{i=1}^{t}\sum_{j=1}^{n}(y_{ij}-\bar{y}_{..})^2=\sum_{i=1}^{t}\sum_{j=1}^{n}(y_{ij}-\bar{y}_{i.})^2+\sum_{i=1}^{t}\sum_{j=1}^{n}(\bar{y}_{i.}-\bar{y}_{..})^2 \qquad (7)$$

وإذا رمزنا لأجمالي مجموع المربعات (التباين الكلي) بالرمز SST ، ولمجموع مربعات المعالجات (بين المعالجات) بالرمز .SSTRT ، ولمجموع مربعات الخطأ (خلال المعالجات) بالرمز SSE فإن :

$$SST = SSTRT. + SSE$$

حيث

$$SST=\sum_{i=1}^{t}\sum_{j=1}^{n}(y_{ij}-\bar{y}_{..})^2=\sum_{i=1}^{t}\sum_{j=1}^{n}y_{ij}^2-\frac{y_{..}^2}{N} \qquad (8)$$

$$SSTRT=\sum_{i=1}^{t}\sum_{j=1}^{n}(\bar{y}_{i.}-\bar{y}_{..})^2=\frac{1}{n}\sum_{i=1}^{t}y_{i.}^2-\frac{y_{..}^2}{N} \qquad (9)$$

اى أن .SSTRT يقيس الاختلاف بيـن المعالجـات (Among treatments) ، بينمـا SSE يقيـس الاختلاف خلال المعالجات أى الاختلاف الناتج عن الخطأ العشوائي (Within treatments) .

كما تم الإشارة سابقاً ، إن الاختبار سيكون مبنى على مقارنة الاختلاف بين المعالجات مع الاختلاف خلال المعالجات وذلك لاكتشاف الفروق المعنوية في المفردات الناتجة عن تأثير المعالجات .وبالنظر الى .SSTRT على أنه متغير عشوائي الذي ستتغير قيمته إذا تمت إعادة التجربة عدة مرات ، فإنه بالإمكان الإثبات بأن :

$$E(SSTRT.) = (t-1)\sigma^2 + n\sum_{i=1}^{t}\alpha_i^2 \qquad (11)$$

اى أن أول تقدير للتباين (σ^2) ، مبنى على (t-1) درجة حرية ويطلق عليه تسمية متوسط مربعات المعالجات وهو معرف كالآتي :

$$MSTRT. = \frac{SSTRT.}{t-1} \qquad (12)$$

وإذا كان فرض العدم صحيحاً ، أى أن جميع $\alpha_i = 0$ فإن :

$$E(MSTRT.) = E\left(\frac{SSTRT.}{t-1}\right) = \sigma^2 \qquad (13)$$

ولكن إذا كان فرض العدم خطأ ،أى أن الفرض البديل صحيحاً فإن :

$$E(MSTRT.) = \sigma^2 + \frac{n\sum_{i=1}^{t}\alpha_i^2}{t-1} \qquad (14)$$

أى أن القيمة المتوقعة لمتوسط مربعات المعالجات تساوي σ^2 مضافاً إليه حد أخـر يقيـس الاختلاف الناتج عن التأثيرات الأخرى .أما التقدير الثاني للتباين المشترك (σ^2) ، والمستقل عن الأول فهو مبنى على t (n-1) درجة حرية ، ويطلق عليه تسمية متوسط مربـع الخطـأ ويرمـز لـه بالرمز MSE حيث :

$$MSE = \frac{SSE}{t(n-1)} \qquad (15)$$

ويمكن الإثبـات بـأن $E(MSE) = \sigma^2$ ، اى أن MSE مقدر غير متحيز للتبـاين المشـترك (σ^2) بغض النظر في ما إذا كان فرض العدم صحيح أم خطأ .إذن يتضح مما سبق أنـه لكـي

يكون .MSTRT مقدراً مقبولاً للتباين المشترك (σ^2) يجب أن يتحقق الشرط الذي ينص على تساوي تباينات جميع المجتمعات وأن يكون فرض العدم صحيحاً ، بينما يكون MSE مقدراً مقبولاً للتباين للمشترك (σ^2) إذا تحقق شرط تساوى تباينات جميع المجتمعات ولكن ليس بالضرورة أن يكون فرض العدم صحيحاً . ولقد تم الإثبات رياضياً أنه عندما يكون فرض العدم صحيحاً يكون التوزيع الاحتمالي للإحصاءة .MSTRT يتبع توزيع مربع كاى وبدرجات حرية تساوى (t - 1) ، بينما التوزيع الاحتمالي للإحصاءة MSE يتبع توزيع مربع كاى وبدرجات حرية تساوى (n - 1) t . وعليه عندما يكون فرض العدم صحيحاً فإن التناسب

$$F = \frac{MSTRT.}{MSE} \qquad (16)$$

يتوزع وفق توزيع f وبدرجات حرية تساوى t-1 و (n -1) t ، وحيث أن .MSTRT يقدر التباين بأكبر مما يجب عندما يكون فرض العدم خطأ ، أى أن $E(MSTRT.) > \sigma^2$ ، وعليه فإنه سيكون هناك اختبار من طرف واحد فقط ، أى أننا نرفض H_o عندما تكون

$$F > f_{\alpha,\, t-1,\, t(n-1)}$$

وعادةً ما يتم وضع الاختلاف الكلى في البيانات في جدول يطلق عليه تسمية جدول تحليل التباين (ANOVA) وذلك على النحو التالي :

جدول تحليل التباين الأحادي (ANOVA)

مصدر الاختلاف S. v.	مجموع المربعات SS	درجات الحرية df	متوسط الخطأ MS	F المحسوبة
المعالجات .TRT	SSRTR	t - 1	$MSTRT. = \frac{SSTRT.}{t-1}$	$F = \frac{MSTRT.}{MSE}$
الخطأ Error	SSE	n (t - 1)	$MSE = \frac{SSE}{n(t-1)}$	
المجموع Total	SST	nt - 1		

لاحظ أنه قد لا يكون عدد المشاهدات في جميع المعالجات متساوي وهذا لا يعنى أن التحليل السابق غير صحيح فهو لا يزال صحيحاً ولكن مع تغيير بسيط في صيغ مجموع المربعات، فإذا

كانت n_i تمثل عدد المفردات في المعالجة i حيث $i = 1,2,\dots,t$ و $N = \sum_{j=1}^{t} n_j$ فإن مجموع المربعات الإجمالي يكون كالآتي :

$$SST = \sum_{i=1}^{t}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{..})^2 = \sum_{i=1}^{t}\sum_{j=1}^{n_i} y_{ij}^2 - \frac{y_{..}^2}{N} \qquad (17)$$

ومجموع مربعات المعالجات (بين المعالجات) يساوي :

$$SSTRT. = \sum_{i=1}^{t}\sum_{j=1}^{n_i}(\bar{y}_{i.} - \bar{y}_{..})^2 = \sum_{i=1}^{t} n_i(\bar{y}_{i.} - \bar{y}_{..})^2 = \sum_{i=1}^{t}\frac{y_{i.}^2}{n_i} - \frac{y_{..}^2}{N} \qquad (18)$$

ومجموع مربعات الخطأ (خلال المعالجات) يساوي :

$$SSE = \sum_{i=1}^{t}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{i.})^2 = \sum_{i=1}^{t}\sum_{j=1}^{n_i} y_{ij}^2 - \sum_{i=1}^{t}\frac{y_{i.}^2}{n_i} = SST - SSTRT. \qquad (19)$$

مثال (1) : بفرض أن ثلاث مجموعات من الطلاب يدرسون نفس المادة في الإحصاء تم تدريسهم من قبل ثلاث أعضاء هيئة تدريس ، فكانت درجات طلاب المجموعات الثلاثة في نهاية الفصل الدراسي كما يلي :

A	B	C
95	85	79
32	90	92
47	79	63
75	50	68
83	32	76
84	84	20
73	78	37
68	95	74
	65	86
	80	
$y_{1.} = 557$	$y_{2.} = 738$	$y_{3.} = 595$
$\bar{y}_{1.} = 69.6$	$\bar{y}_{2.} = 73.8$	$\bar{y}_{3.} = 66.1$
$n_1 = 8$	$n_2 = 10$	$n_3 = 9$

من هذه البيانات هل يمكن القول بأن هناك فروق معنوية بين متوسطات الدرجات المعطاة من قبل الأساتذة الثلاث عند مستوى المعنوية 5 ٪ ؟

الحـل :

الفرضية :

إن الفرضية المطلوب اختبارها تكون على النحو الآتي :

$$H_o : \mu_1 = \mu_2 = \mu_3$$

H_1 : على الأقل متوسطين مختلفين

إحصاءة الاختبار :

يمكن اختبار الفرضية أعلاه كما يلي :

من الجدول أعلاه نجد أن $y_{..} = 1890$ ، وإن $\bar{y}_{..} = \frac{1890}{27} = 70$

$$SST = \sum_{i=1}^{3}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{..})^2 = (95-70)^2 + (32-70)^2 + \cdots + (86-70)^2 = 11076$$

$$SSTRT. = \sum_{i=1}^{3} n_i\ (\bar{y}_{i.} - \bar{y}_{..})^2 = 8(69.6-70)^2 + 10(73.8-70)^2$$

$$+9(66.1-70)^2 = 281.64$$

$$SSE = \sum_{i=1}^{3}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{i.})^2 = (95-69.6)^2 + (32-69.6)^2 + \ldots + (68-69.6)^2$$

$$+ (85-73.8)^2 + \cdots + (80-73.8)^2$$

$$+ \cdots + (79-66.1)^2 + \ldots + (86-66.1)^2 = 10794.36$$

ألحظ أنه يمكن الحصول على القيم أعلاه بسهولة أكثر باستخدام الصيغ العملية (كما سنرى في المثال القادم)

وبالتالي يكون جدول تحليل التباين (ANOVA) لهذا المثال كما يلي :

مصدر الاختلاف	مجموع المربعات	درجات الحرية	متوسط المربع	F المحسوبة
المجموعات .TRT	281.64	2	140.82	0.3131
Error الخطأ	10794.36	24	449.765	
Total المجموع الكلي	11076.0	26		

القرار :

من جدول F وبدرجات حرية تساوى 2 و 24 و $\alpha = 0.05$ نجد ان $f_{0.05,2,24} = 3.4$ ، وحيث ان

$0.3131 < 3.4$ ، وعليه لا توجد معلومات كافية لرفض H_0 .

مثال (2) : اجريت تجربة في مركز الأبحاث الحيوانية على أربعة أنواع مختلفة من الفيتامينات يعتقد أنها تسبب في زيادة أوزان الأبقار ، فإذا تم اختيار 20 بقرة وتم توزيعها بطريقة عشوائية على الفيتامينات الأربعة وبعد فترة زمنية معينة كانت نتائج الزيادة في الوزن كما يلي :

A	B	C	D
46	65	27	15
53	59	37	30
54	43	25	27
39	49	35	28
33	37	43	40
$y_{i.} = 225$	253	167	140
$\bar{y}_{i} = 45$	50.6	33.4	28

من هذه البيانات هل يمكن القول بأن هناك فروق جوهرية ما بين الفيتامينات الأربعة عند مستوى

المعنوية 5 % ؟

الحل :

لتحليل هذه البيانات سوف نفترض على أنها تمثل عينة عشوائية من أربع مجتمعات متشابهة عدا في نوعية الفيتامين الذي تم استخدامه وإن لهذه المجتمعات توزيع طبيعي وبتباين متساوي .

الفرضية :

$$H_0 : \mu_A = \mu_B = \mu_C = \mu_D$$

H_1 : على الأقل أحدها يختلف عن بقية المتوسطات

و

$$H_0 : \alpha_i = 0$$
$$H_1 : \alpha_i \neq 0$$

إحصاءة الاختبار :

حيث أن

$$\sum_{i=1}^{t}\sum_{j=1}^{n} y_{ij}^2 = (46)^2 + (53)^2 + \ldots + (28)^2 + (40)^2 = 33811$$

$$\frac{1}{n}\sum_{i=1}^{t} y_{i.}^2 = \frac{1}{5}\left[(225)^2 + (253)^2 + (167)^2 + (140)^2\right] = \frac{162123}{5} = 32424.6$$

$$y_{..} = \sum_{i=1}^{t}\sum_{j=1}^{n} y_{ij} = 46 + 53 + \cdots + 28 + 40 = 785$$

$$\Rightarrow \frac{y_{..}^2}{tn} = \frac{(785)^2}{20} = \frac{616225}{20} = 30811.25$$

$$SSTRT = \frac{1}{n}\sum_{i=1}^{4} y_{i.}^2 - \frac{y_{..}^2}{tn}$$
$$= 32424.6 - 30811.25 = 1613.35$$

$$SSE = \sum_{i=1}^{4}\sum_{j=1}^{5} y_{ij}^2 - \frac{1}{n}\sum_{i=1}^{4} y_{i.}^2$$
$$= 33811 - 32424.6 = 1386.4$$

$$SST = \sum_{i=1}^{4}\sum_{j=1}^{5} y_{ij}^2 - \frac{y_{..}^2}{tn}$$
$$= 33811 - 30811.25 = 2999.75$$

وعليه فإن :

$$MSTRT = \frac{SSTRT}{t-1} = \frac{1613.35}{3} = 537.7833$$

$$MSE = \frac{SSE}{t(n-1)} = \frac{1386.4}{4 \times 4} = \frac{1386.4}{16} = 86.65$$

ويمكن وضع هذه النتائج في جدول تحليل التباين (ANOVA) وذلك كما يلي :

المصدر	درجات حرية	مجموع المربعات	متوسط المربعات	F المحسوبة
المعالجات TRT	3	1613.35	537.7833	$\frac{537.7833}{68.65} = 6.21$
الخطأ Error	16	1386.4	68.65	
المجموع Total	19	2999.75		

القرار :

من جدول F وبدرجات حرية تساوي 3 و 16 ومستوى معنوية $\alpha = 0.05$ نجد أن : $f_{0.05,3,16} = 3.24$ ، وحيث أن 6.21 أكبر من 3.24 ، وعليه نرفض H_0 عند مستوى المعنوية 5 % . وبالتالي يمكن القول بأنه ليس للفيتامينات الأربعة نفس التأثير من حيث الزيادة في الوزن عند هذا المستوى من المعنوية .

ملحوظة :

1 - إن نوع التجربة الذي تم تغطيته يطلق عليه تسمية تجربة بتأثير ثابت (fixed effect) ، وذلك لأن ما يهمنا في تحليل التباين هنا هو المعالجات فقط .

2 - عندما يكون عدد المعالجات يساوى 2 ($H_0 : \mu_1 = \mu_2$) فإنه يمكن استخدام اختبار t الذي سبق وأن تعرضنا إليه في فصل اختبارات الفرضيات الإحصائية ، وإنه يمكن الإثبات في هذه الحالة بأن

$F = \frac{MSTRT}{MSE} = t^2$ وإن $t^2_{\frac{\alpha}{2}, N-a} = f_{\alpha,1,N-a}$ ، حيث a تعني عدد المعالجات هنا بينما t تعني اختبار t المألوف .

3 - إن MSE يقدر σ^2 وإن $s_i^2 = \frac{\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{i.})^2}{n_i - 1}$ يقدر تباين المجموعة (المعالجة) i ، ويمكن الإثبات بأن $MSE = \frac{\sum_{i=1}^{t}(n_i - 1)s_i^2}{N - t}$.

4 - إن هذا النوع من التجربة يطلق عليه تسمية التصميم العشوائي الكامل ، وذلك لأن الوحدات التجريبية (N) تم تصنيفها أو تخصيصها للمعالجات (t) بطريقة عشوائية .

10 - 3 المقارنات المتعددة multiple comparisons

لقد أشرنا فيما سبق أن الهدف من تحليل التباين هو اختبار الفرضيات الخاصة بالمتوسطات وذلك من أجل تحديد الفروق إن وجدت بين هذه المتوسطات تم تقدير تلك الفروق إن كانت معنوية .

ومن خلال العرض السابق لتحليل التباين تبين لنا أنه لاختبار الفرضيات الخاصة بمتوسطات المعالجات أو تأثيرها نكون جدول تحليل التباين ثم نستخدم اختبار F ، ومن خلال نتيجة هذا الاختبار يتبين فيما إذا كان هناك فروق بين المتوسطات أم لا فإذا كانت خلاصة الاختبار تشير إلى وجود فروق جوهرية بين متوسطات المعالجات فإننا نقول بأن الاختبار معنوي والسؤال الذي يطرح نفسه بعد ذلك هو : بين أي المتوسطات توجد تلك الفروق أو الاختلافات ؟

وللإجابة على هذا التساؤل يتطلب الأمر إجراء عدة مقارنات بين متوسطات المعالجات لتحديد أي المتوسطات تختلف عن بعضها البعض . إن مثل هذا الإجراء يطلق عليه تسمية المقارنات المتعددة . وهناك العديد من الطرائق المقترحة لهذا الغرض وسوف نتعرض لبعض منها في هذا البند .

أ - المقارنات المصممة الخطية Linear Contrasts

في بعض التجارب يرغب الباحث في إجراء مقارنات بين متوسطات معالجات محددة قبل تنفيذ التجربة أي في هيئة مقارنات مصممة ، فمثلاً وإذا كان هناك تجربة ما وتتضمن 4 معالجات فمن الممكن أن الباحث يرغب في مقارنة المعالجتين " 1 و 2 " مع المعالجتين " 3 و 4 "، أو مقارنة المعالجات " 1 و 2 و 3 " مع المعالجة 4 ... وهكذا .

تعريف : المقارنات المصممة (linear contrasts) أو المحددة سلفاً هي عبارة عن علاقة خطية في متوسطات المعالجات ، فإذا كانت $\mu_1, \mu_2, \mu_3, \dots, \mu_t$ تمثل متوسطات " t " من المعالجات فإن العلاقة الخطية في هذه المتوسطات تكون على النحو الآتي :

$$L = \sum_{i=1}^{t} c_i \mu_i \qquad (20)$$

بحيث $\sum_{i=1}^{t} c_i = 0$ ، فمثلاً إذا كانت $\mu_1, \mu_2, \mu_3, \mu_4$ تمثل متوسطات أربعة معالجات فإنه من الممكن أن تكون المقارنات المصممة الخطية كما يلي :

1- مقارنة متوسط المعالجة الأولى مع الثانية : $L = \mu_1 - \mu_2$

$$\sum_{i=1}^{4} c_i = 0 \quad , c_1 = 1 \, , c_2 = -1 \quad , \quad c_3 = c_4 = 0$$ وعليه فإن

2- مقارنة متوسطات المعالجتين " 1 و 2 " مع متوسطات المعالجتين " 3 و 4 " :

$$L = \left(\frac{\mu_1 + \mu_2}{2}\right) - \left(\frac{\mu_3 + \mu_4}{2}\right)$$

إذن $c_3 = c_4 = -\frac{1}{2}$ ، $c_1 = c_2 = \frac{1}{2}$ وعليه فإن $\sum_{i=1}^{4} c_i = 0$.

وهكذا يمكن كتابة أى صيغة خطية في متوسطات المعالجات . إذن إذا كانت L تمثل علاقة خطية في متوسطات المعالجات فإنه يمكن اختبار الفرضية الآتية :

$$H_0 : L = 0$$
$$H_1 : L \neq 0$$

وذلك باستخدام اختبار t المألوف على النحو الآتي :

$$t = \frac{\hat{L}}{\sqrt{(MSE)\sum_{i=1}^{t} \frac{c_i^2}{n_i}}} \qquad (21)$$

حيث $\hat{L} = \sum_{i=1}^{t} c_i \bar{y}_{i.}$ و MSE يمثل متوسط مربع الخطأ بجدول تحليل التباين ، وسوف نرفض H_0 عند مستوى المعنوية α إذا كانت $t \geq t_{\frac{\alpha}{2}, N-t}$ أو $t \leq -t_{\frac{\alpha}{2}, N-t}$.

وبالمثل يمكن اختبار الفرضية إذا كانت من طرف واحد .أيضاً يمكن إيجاد $100\%(1-\alpha)$. فترة ثقة حول L وذلك كما يلي :

$$\hat{L} - (t_{\frac{\alpha}{2}, N-t}) \sqrt{MSE \sum_{i=1}^{t} \frac{c_i^2}{n_i}} \leq L \leq \hat{L} + (t_{\frac{\alpha}{2}, N-t}) \sqrt{MSE \sum_{i=1}^{t} \frac{c_i^2}{n_i}} \qquad (22)$$

مثال (3) : استخدام بيانات المثال (2) لاختبار الفرضية الآتية :

$$H_0 : \frac{1}{2}(\mu_A + \mu_B) - \frac{1}{2}(\mu_C + \mu_D) = 0$$
$$H_1 : \frac{1}{2}(\mu_A + \mu_B) - \frac{1}{2}(\mu_C + \mu_D) \neq 0$$

في الفرضية أعلاه يمكن إعادة كتابتها كما يلي :

$$H_0 : L = 0$$
$$H_1 : L \neq 0$$
$$L = \frac{1}{2}\mu_1 + \frac{1}{2}\mu_2 - \frac{1}{2}\mu_3 - \frac{1}{2}\mu_4$$
$$= \sum_{i=1}^{4} c_i \mu_i \quad , c_1 = c_2 = \frac{1}{2} \quad , c_3 = c_4 = -\frac{1}{2}$$

حيث $\mu_4 = \mu_D$ و $\mu_3 = \mu_C$ و $\mu_2 = \mu_B$ و $\mu_1 = \mu_A$.ويتم تقدير L كما يلي :

$$\hat{L} = \sum_{i=1}^{4} c_i \hat{\mu}_i = \sum_{i=1}^{4} c_i \bar{y}_i$$
$$= \frac{1}{2}\bar{y}_1 + \frac{1}{2}\bar{y}_2 - \frac{1}{2}\bar{y}_3 - \frac{1}{2}\bar{y}_4$$
$$= \frac{1}{2}(45) + \frac{1}{2}(50.6) - \frac{1}{2}(33.4) - \frac{1}{2}(28) = 17.1$$

وأن

$$\sum_{i=1}^{4} \frac{c_i^2}{n_i} = \frac{1}{5}\left[\left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2 + \left(-\frac{1}{2}\right)^2 + \left(-\frac{1}{2}\right)^2\right]$$
$$= \frac{1}{5}[1] = \frac{1}{5} = 0.2$$
$$\sqrt{(MSE)\sum_{i=1}^{4} \frac{c_i^2}{n_i}} = \sqrt{(86.65)(0.2)} = 4.1629$$
$$t = \frac{\hat{L}}{\sqrt{MSE\sum_{i=1}^{4} \frac{c_i^2}{n_i}}} = \frac{17.1}{4.1629} = 4.11$$

ومن جدول t وبدرجات حرية تساوى $N - t = 16$ و $\alpha = 0.05$ نجد أن :

$t_{\frac{\alpha}{2}, N-t} = t_{0.025, 16} = 2.121$ ، وحيث أن 4.11 أكبر من 2.12 ، وعليه نرفض H_0 ويمكن القول بأن تأثير A و B يختلف على تأثير C و D .

إن تطبيقات هذا الأسلوب محدودة وذلك لأن معامل الثقة $(1-\alpha)$ أو مستوى المعنوية (α) يُطبق على تقدير (أو اختبار) معين فقط ، وبالتالي لا يمكن تطبيقه أيضاً على سلسلة من

التقديرات أو الاختبارات ، وذلك لأن هذا المعامل أو المستوى مـن الثقـة يكـون مناسباً للتقدير أو الاختبار غير المقترح من قبل البيانات وبالتالي دعت الحاجة إلي استخدام طرائق أخرى سنتناول بعض منها في هذا البند .

ب - اختبار توكى (Tukey's honestly significant difference (HSD

إن أسلوب المقارنات المتعددة المقترح من قبل توكي عادةً ما يستخدم لاختبار الفرضيات التـى تتضمن أن جميع الأزواج الممكنة لمتوسطات المعالجات متساوية أى أن :

$$H_0 : \mu_i = \mu_j$$
$$H_1 : \mu_i \neq \mu_j$$

وعند تطبيق هذا الاختبار يتم اختيار مستوى معنويـة وحجم العينة في جميع المعالجات متساوي ، وإن هذا عام (α) وبالتالي فإن احتمال أن يكون فرض عدم أو أكثر غير صحيـح يسـاوى α ، الاختبار يستخدم قيمة حرجة واحدة فقط مقابل جميع الفروق المقارنة هذه القيمة يرمز لها بالرمز HSD ومعرفة كما يلي :

$$HSD = q_{\alpha, t, N-t}\sqrt{\frac{MSE}{n}} \qquad (23)$$

حيث " t " تمثل عدد المعالجات و N العـدد الكلـى للمفـردات و n تمثـل عـدد المفـردات فـى كـل معالجة و MSE تمثل متوسط مربع الخطأ فى جدول تحليـل التبـاين و q تمثـل القيمـة الجدوليـة بجدول (16) وسوف نرفض H_0 إذا كانت : $\left|\bar{y}_{i.} - \bar{y}_{j.}\right| > HSD$.

الحظ أنه إذا كان حجم العينة غير متساوى فى جميع المعالجات فإنـه يتـم إسـتبدال n فـى الصيغـة أعلاه بالآتى :

$$\tilde{n} = \frac{t}{\frac{1}{n_1} + \frac{1}{n_2} + \ldots + \frac{1}{n_t}}$$

حيث t تمثل عدد المعالجات ويمكن تلخيص النتائج من خلال وضع خط تحت المتوسطات التى لا توجد بينها فروق ، أما إذا كانت الفروق معنوية بين أى متوسطين فلا نضع تحتها خط وسـنوضح ذلك من خلال المثال القادم .

مثال (4) : استخدام اختبار HSD لتحديد أى المتوسطات تختلف عن بعضها البعض في مثال (2) .

الحل :

من نتائج ذلك المثال نحن نعلم أن

$$\bar{y}_D = 28 \;,\; \bar{y}_C = 33.4 \;,\; \bar{y}_B = 50.6 \;,\; \bar{y}_A = 45$$

وعليه فإن ترتيب هذه المتوسطات يكون كما يلي :

البنامين	A	B	C	D
ترتيب المتوسطات	2	1	3	4

وعليه فإن جميع الفروق المطلقة الممكنة (مرتبة) ما بين المتوسطات يمكن تلخيصها في الجدول التالي :

	B	A	C	D
B	–	5.6	17.2	22.6
A	–	–	11.6	17.0
C	–	–	–	5.4
D	–	–	–	–

حيث أن $\alpha = 0.05$ و $N\text{-}t = 16$ و $MSE = 86.65$ وبالتالي من جدول (16) نجد أن $q_{\alpha,t,N-t} = q_{0.05,4,16} = 4.05$ وعليه فإن

$$HSD = q_{\alpha,t,N-t}\sqrt{\frac{MSE}{n}} = (4.05)\sqrt{\frac{86.65}{5}} = 16.86$$

وبالتالي فإن :

الفرضية	الفرق المطلق ما بين المتوسطين	HSD	القرار
$H_0:\mu_A=\mu_B$	5.6	16.86	لا نرفض H_0 (غير معنوي)
$H_0:\mu_B=\mu_C$	17.2	16.80	نرفض H_0 (معنوي)
$H_0:\mu_B=\mu_D$	22.6	16.86	نرفض H_0 (معنوي)
$H_0:\mu_A=\mu_C$	11.6	16.86	لا نرفض H_0 (غير معنوي)
$H_0:\mu_A=\mu_D$	17	16.86	نرفض H_0 (معنوي)
$H_0:\mu_C=\mu_D$	5.4	16.86	لا نرفض H_0 (غير معنوي)

إن نتائج الجدول السابق يمكن تلخيصها بالطريقة المقترحة من قبل دنكن (Duncan) وذلك بوضع خط يربط ما بين أى متوسطين أو أكثر لا يختلفان .

$$\underline{\mu_B \quad \mu_A} \quad \mu_C \quad \mu_D$$

وعليه يمكن القول بأن μ_B و μ_A هما الأفضل من حيث التأثير على الزيادة في الوزن بينما μ_D هو الأقل .

جـ - اختبار فيشر لأقل فرق معنوي **Fisher's least significant difference (L.S.D.)**

يستخدم هذا الاختبار للمقارنة بين متوسطي كل معالجتين على حده ويوصى فيشر بعدم استخدام هذا الاختبار إلا إذا كان اختبار F معنوي ، وبالتالي يطلق عليه أحياناً تسمية الاختبار المحفوظ (Protected) . وصيغة أقل فرق معنوي (LSD) تكون كما يلي :

$$L.S.D=(t_{\frac{\alpha}{2},N-t})\sqrt{MSE\left(\frac{1}{n_i}+\frac{1}{n_j}\right)} \qquad (24)$$

حيث t تمثل عدد المعالجات و N تمثل العدد الكلي للمفردات بالتجربة وسوف نرفض $H_0:\mu_i=\mu_j$ إذا كانت :

$$\left|\bar{y}_{i.}-\bar{y}_{j.}\right|>L.SD.$$

وإذا كان حجم العينة متساوي في جميع المعالجات فإنه ستكون هناك قيم واحدة يتم المقارنة بها وهى :

$$L.S.D. = (t_{\frac{\alpha}{2},N-t})\sqrt{\frac{2\ MSE}{n}} \quad (25)$$

مثال (5) : استخدم اختبار فيشر لتحديد أى المتوسطات تختلف عن بعضها البعض في المثال رقم (2) .

الحل :

حيث أن $\alpha=0.05$ و $N-t=16$ وعليه فإن : $t_{\frac{\alpha}{2},N-t} = t_{0.025,16} = 2.12$ وحيث أن حجم العينة متساوي في جميع المعالجات وبالتالي فإن قيمة .L.S.D ستكون متساوية في جميع المقارنات ، وبالتالي فإن :

$$L.S.D. = (t_{\frac{\alpha}{2},N-t})\sqrt{\frac{2\ MSE}{n}} = (2.12)\sqrt{\frac{2(86.65)}{5}} = 12.4805$$

وبنفس الطريقة التى أتبعناها في حالة اختبار .H.S.D من حيث إيجاد جميع الفروق المطلقة الممكنة ما بين المتوسط سيكون اختبار L.S.D كما يلي :

الفرضية	الفرق المطلق ما بين المتوسطين	HSD	القرار
$H_0:\mu_A=\mu_B$	5.6	12.4805	لا نرفض H_0 (غير معنوي)
$H_0:\mu_B=\mu_C$	17.2	12.4805	نرفض H_0 (معنوي)
$H_0:\mu_B=\mu_D$	22.6	12.4805	نرفض H_0 (معنوي)
$H_0:\mu_A=\mu_C$	11.6	12.4805	لا نرفض H_0 (غير معنوي)
$H_0:\mu_A=\mu_D$	17	12.4805	نرفض H_0 (معنوي)
$H_0:\mu_C=\mu_D$	5.4	12.4805	لا نرفض H_0 (غير معنوي)

وعليه فإن μ_D μ_C μ_A μ_B وهى نفس النتيجة التى توصلنا إليها عند استخدام اختبار HSD .

د - اختبار دونت **Dunnett's test**

في بعض التجارب يهدف الباحث الى مقارنة مجموعة من المعالجات مع معالجة يطلق عليها تسمية معالجة السيطرة أو المراقبة (Placebo or control treatment) ، وذلك عوضاً عن مقارنات المعالجات مع بعضها البعض كما في الحالات السابقة، أى أن الفرضية التى يرغب في اختبارها تكون على النحو الآتي :

$$\left.\begin{array}{l} H_o : \mu_0 = \mu_i \\ H_1 : \mu_0 \neq \mu_i \end{array}\right\} , i = 1, 2, \cdots, t$$

حيث μ_0 تمثل متوسط مجمتع معالجة المراقبة و μ_i تمثل متوسط مجتمع المعالجة i ، ويمكن تطبيق هذا الاختبار حتى إذا كان اختبار F غير معنوي ، وذلك لأن هذه المقارنات سبق وأن حددها الباحث قبل تنفيذ التجربة وليست مقترحة من البيانات بعد إجراء التجربة عليها ، ولتطبيق هذا الاختبار يفترض أن شروط تحليل التباين التى سبق وأن أشرنا إليها تكون محققة .ولاختبار هذه الفرضية توجد أولاً القيم :

$$d_i = \frac{\bar{y}_{i.} - \bar{y}_0}{\sqrt{\frac{2\,MSE}{n}}} \quad , \; i = 1, 2, \cdots, t \qquad (26)$$

حيث n تمثل عدد المشاهدات بكل معالجة ، وMSE يمثل متوسط مربع الخطأ بجدول تحليل التباين ، وسوف نرفض H_0 إذا كانت $|d_i| > d_{\frac{\alpha}{2}, t-1, N-t}$ حيث $d_{\frac{\alpha}{2}, t-1, N-t}$ تمثل القيمة الجدولية لاختبار دونت (جدول 15) في حالة ما يكون الاختبار من طرفين ، أم إذا كان الاختبار من طرف واحد أى أن

$$\left.\begin{array}{l} H_0 : \mu_i = \mu_0 \\ H_1 : \mu_i > \mu_0 \end{array}\right\} , i = 1, 2, \cdots, t$$

فإننا نرفض $_0$ كانت $|d_i| > d_{\alpha, t-1, N-t}$.

مثال (6) : باحث في علم الأحياء استخدم 4 مستويات من مركب كيميائي معين وذلك لغرض معرفة تأثير هذه المستويات من المركب على الزيادة في أطوال نوع معين من النبات بعد فترة زمنية معينة من تطبيق هذا المركب ولقد أستخدم 5 نباتات لكل مستوى وتم قياس الطول لكل

نبات بعد الفترة الزمنية المحددة للتجربة واستخدم نبات كمراقب أي أنه لم يتم معالجته بالمركب المستعمل بالتجربة فكانت النتائج كما يلي :

التركيز				
0	1	2	3	4
5.9	8.2	7.7	6.9	6.8
6.1	8.7	8.4	7.3	7.3
6.9	9.4	8.6	6.3	6.3
5.7	9.2	8.1	6.8	6.9
6.1	8.6	8.0	7.4	7.1

عند مستوى المعنوية 5٪ هل يمكن القول بوجود فروق معنوية بين كل مستوى من مستويات التركيز ومعالجة المراقبة .

الحل :

من الجدول اعلاه نجد أن

$$y_{0} = 30.70 \ , \ y_{1} = 44.1 \ , \ y_{2} = 40.8 \ , \ y_{3} = 34.1 \ , \ y_{4} = 34.4$$

$$\bar{y}_{0} = 6.14 \ , \ \bar{y}_{1} = 8.82 \ , \ \bar{y}_{2} = 8.16 \ , \ \bar{y}_{3} = 6.82 \ , \ \bar{y}_{4} = 6.88$$

وإن

$$\sum_{i=0}^{4} y_{i}^{2} = (30.70)^2 + (44.10)^2 + (40.80)^2 + (34.10)^2 + (34.40)^2 = 6898.11$$

$$\sum_{i=0}^{4} \sum_{j=1}^{5} y_{ij}^{2} = (5.9)^2 + (6.1)^2 + \cdots + (6.9)^2 + (7.1)^2 = 1384.0$$

وعليه فإن :

$$SSE = \sum_{i=0}^{4} \sum_{j=1}^{5} y_{ij}^{2} - \left(\frac{\sum_{i=0}^{4} y_{i}^{2}}{n} \right) = 1384.0 - \frac{6898.11}{5} = 4.378$$

$$\Rightarrow MSE = \frac{4.378}{20} = 0.2189 \Rightarrow \sqrt{\frac{2MSE}{n}} = \sqrt{\frac{2(0.2189)}{5}} = 0.2959$$

ومن جدول (15) حيث أن $\alpha = 0.05$ و $t = 4$ و $N-t = 20$ نجد أن :

$$d_{\frac{\alpha}{2}, t-1, N-t} = d_{0.025, 4, 20} = 2.65$$

وعليه فإن :

الفرضية	إحصاءة الاختبار	القيمة الجدولية	القرار :نرفض H_0 إذا كان $\|d_i\| > 2.65$
$H_o:\mu_0 = \mu_1$ $H_1: \mu_0 \neq \mu_1$	$d_1 = \frac{8.82 - 6.14}{0.2959} = 9.0571$	2.65	نرفض H_0 (معنوي)
$H_o:\mu_0 = \mu_2$ $H_1: \mu_0 \neq \mu_2$	$d_2 = \frac{8.16 - 6.14}{0.2959} = 6.8266$	2.65	نرفض H_0 (معنوي)
$H_o:\mu_0 = \mu_3$ $H_1: \mu_0 \neq \mu_3$	$d_3 = \frac{6.82 - 6.14}{0.2959} = 2.2981$	2.65	لا نرفض H_0 (غير معنوي)
$H_o:\mu_0 = \mu_4$ $H_1: \mu_0 \neq \mu_4$	$d_4 = \frac{6.68 - 6.14}{0.2959} = 1.8249$	2.65	لا نرفض H_0 (غير معنوي

ومن الجدول أعلاه يتضح أن لتركزين الأول والثاني تعطى فروق معنوية عن متوسط طول النبات بدون إضافة التركيز .

10 - 4 اختبار تساوى عدة تباينات Test for the equality of several variances

مما سبق يتضح ان احد الشروط المطلوبة لصحة تحليل التباين هو أن يكون لكل مجتمع من المجتمعات قيد الدراسة مجتمع طبيعي وأن تكون العينات المتحصل عليها من هذه المجتمعات عشوائية ومستقلة وبعبارة أخرى يجب أن تكون الأخطاء العشوائية متغيرات عشوائية مستقلة ولها توزيع طبيعي بمتوسط يساوى صفر وتباين يساوى σ^2 أي أن التباين في جميع المجتمعات متساوي بالرغم من أن اختبار F بتحليل التباين غير حساس لهذا الشرط بدرجة كبيرة عندما يكون عدد المشاهدات في جميع المجتمعات (المعالجات) متساوي ، ومع ذلك يفضل إجراء اختبار لمعرفة مدى تجانس التباينات وخاصة عندما يكون عدد المشاهدات غير متساوي في جميع المعالجات .أي أن الفرضية المطلوب التحقق من صحتها تكون على النحو الآتي :

$$H_o:\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \cdots = \sigma_t^2$$

$H_1:$ ليست جميع التباينات متساوية

إن أحد الاختبارات المقترحة لاختبار تجانس تباينات عدة مجتمعات هو اختبار بارتليت (Bartlett test) .إن هذا الاختبار يعطى قيم حرجة ودقيقة عندما يكون عدد المشاهدات متساوي

جميع العينات ، ولكن سوف لن تكون هذه القيم دقيقة بشكل كبير عندما يكون عدد المشاهدات في جميع العينات غير متساوي ومع ذلك لا زال الاختبار جيد إلى حد كبير ، ولتطبيق هذا الاختبار نتبع الآتي :

نوجد أولاً تباينات جميع العينات ففي حالة تحليل التباين الأحادي تكون

$$s_i^2 = \frac{\sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{i.})^2}{n_i - 1} \quad , \ i = 1, 2, \cdots, t . \tag{27}$$

ثم نوجد التباين المشترك لهذه التباينات ، أي $s_p^2 = \frac{\sum_{i=1}^{t} (n_i - 1) s_i^2}{N - t}$ حيث $N = \sum_{i=1}^{t} n_i$.

أخيراً إحصاءة الاختبار تكون كالآتي :

$$B = \frac{w}{c} \tag{28}$$

حيث

$$w = (N - t) \ln s_p^2 - \sum_{i=1}^{t} (n_i - 1) \ln s_i^2 \tag{29}$$

و

$$c = 1 + \frac{1}{3(t-1)} \{ \sum_{i=1}^{t} \left(\frac{1}{n_i - 1}\right) - \left(\frac{1}{N - t}\right) \} \tag{30}$$

وسوف نرفض H_0 إذا كانت $B > \chi^2_{\alpha, t-1}$ حيث $\chi^2_{\alpha, t-1}$ تمثل قيمة مربع كاي الجدولية .

مثال (7): بفرض أن البيانات التالية تمثل نتائج ثلاث عينات تم الحصول عليها من ثلاثة مجتمعات :

العينة	1	2	3
حجم العينة n_i	20	17	21
تباين العينة s_i^2	415	698	384

من هذه البيانات هل يمكن القول بأن تباينات مجتمعات هذه العينات متجانسة عند مستوى المعنوية 5% ؟

الحـل :

الفرضية :

بفرض أن σ_1^2 و σ_2^2 و σ_3^2 تمثل تباين المجتمع الأول و الثاني والثالث على التوالي ،

وبالتالي فإن :

$$H_0 : \sigma_1^2 = \sigma_2^2 = \sigma_3^2$$

H_1 : التباينات غير متجانسة

إحصاءة الاختبار :

الحسابات الضرورية لحساب الاحصاءة موضحة في الجدول التالي :

المجتمع	$n_i - 1$	$(n_i - 1) s_i^2$	$\ln s_i^2$	$(n_i - 1) \ln s_i^2$
1	19	7885	6.02829	114.53751
2	16	11168	6.54822	104.77152
3	20	7680	5.95064	119.0128
المجموع	55	26733		338.32183

وإن $s_p^2 = \dfrac{\sum_{i=1}^{t}(n_i - 1)s_i^2}{N - t} = \dfrac{26733}{55} = 486.05$ و $\ln s_p^2 = 6.18631$ و

$$c = 1 + \frac{1}{3(t-1)}\left[\sum_{i=1}^{t}\left(\frac{1}{n_i - 1}\right) - \left(\frac{1}{N-1}\right)\right]$$

$$= 1 + \frac{1}{3(3-1)}\left[\left(\frac{1}{19} + \frac{1}{16} + \frac{1}{20}\right) - \frac{1}{55}\right] = 1.02449$$

$$w = (N - t)\ln s_p^2 - \sum_{i=1}^{t}(n_i - 1)\ln s_i^2$$

$$= (58 - 3)(6.18631) - 338.32183 = 1.9253$$

وعليه فإن :

$$B = \frac{w}{c} = \frac{1.9253}{1.02449} = 1.8793$$

قرار :

من جدول مربع كاي نجد أن $\chi^2_{\alpha,t-1} = \chi_{0.05,2} = 5.991$ ، وحيث أن $1.8793 < 5.991$ وعليه لا توجد معلومات كافيــة لرفـض H_0 ، أي يمكـن القول بـأن تباينـات المجتمعـات الثـلاث متساوية (متجانسة).

هناك طريقتين يمكن من خلالهما يتم اختيار مستويات العامل لتجربة ما ، الأولى هي أن المعالجات " t " يتم اختيارها بشكل محدد من قبل الباحث ، وذلك من أجل تقدير α_i ، أو لاختبار الفرضية المتعلقة بها ، وفي هذه الحالة سوف تكون الخلاصة منطبقة على مستويات العامل محل الدراسة فقط ، ولا يمكن تعميم هذه الخلاصة والاستنتاج على معالجات مشابهة لم تتم دراستها ، وفي هـذه الحالـة يطلـق علـى نمـوذج تحليـل التبـاين تسـمية نمـوذج التـأثيرات الثابتـة (fixed effects model) وهو ما تعرضنا إليه سلفاً. أما الطريقة الثانية فهي من الممكن أن تكون المعالجات " t " عبارة عن عينة عشوائية من مجتمع كبير من المعالجات، وفي هذه الحالة بإمكـان الباحث تعميم الخلاصة والاستنتاج التي يتحصل عليها من هذه العينة على مجتمع المعالجات ككـل بغض النظر فيما إذا كانت من ضمن عينة المعالجات التي أشملها التحليل أم لا، وبالتـالي يطلـق على هذا النموذج تسمية نموذج التأثيرات المتغيرة (random effects model) ، ويقـوم هـذا النموذج على الافتراضات التالية :

أ - إن مجتمع مستويات العامل إما أن يكـون لانهـائي أو كبـير بشكل كـافي حتـى يمكـن اعتبـاره لانهائي .

ب - إن " α_i " و i=1,2,...,n متغيرات عشـوائية مستقلة ولكـل منهـا توزيـع طبيعـي ، أي أن $\alpha_i \sim N(0,\sigma^2_\alpha)$.

جـ - إن ϵ_{ij} متغيرات عشوائية مستقلة ومتطابقة التوزيع و $\epsilon_{ij} \sim N(0,\sigma^2)$.

د - α_i و ϵ_{ij} مستقلة عن بعضها البعض .

هـ - الفرضية ستكون كالآتي :

$$H_0: \sigma^2_\alpha = 0$$
$$H_1: \sigma^2_\alpha > 0$$

نلاحظ أن شرط أن تكـون α_i متغـيرات عشـوائية ومسـتقلة يعنـي أن الشـرط المـألوف وهـو $\sum_{i=1}^{t} \alpha_i = 0$ في حالة نموذج التأثيرات الثابتة لا ينطبق في حالة نموذج التأثيرات المتغيرة.

ولاختبار الفرضية أعلاه يتطلب الأمر كما في حالة نموذج التأثيرات الثابتة تجزئة إجمالي مجموع المربعات إلى مجموعتين إحداهما يقيس الاختلاف بين المعالجات (SSTRT.) ، والآخر يقس الاختلاف خلال المعالجات (SSE) ، وعليه إذا كانت $\sigma_\alpha^2 = 0$ فإن جميع المعالجات متساوية ، أما إذا كانت $\sigma_\alpha^2 > 0$ فهذا يعني وجود اختلاف ما بين المعالجات ، ويمكن إيجاد القيمة المتوقعة لمتوسط مجموع مربعات المعالجات ومتوسط مجموع مربعات الخطأ ، إلا أننا سوف لن نتعرض لكيفية الاشتقاق هنا بل نكتفي بكتابة صيغة هذه القيم وهي :

$$E(MSTRT.) = E(\frac{SSTRT.}{t-1}) = \sigma^2 + n\sigma_\alpha^2 \qquad (31)$$

حيث افترضنا أن عدد المفردات (n) في كل المعالجات متساوي ، وإن

$$E(MSE) = E(\frac{SSE.}{N-t}) = \sigma^2 \qquad (32)$$

وعليه يتضح من القيم المتوقعة أنه إذا كانت الفرضية $H_o : \sigma_\alpha^2 = 0$ صحيحة فإن كلا من E (MSTRT.) و E(MSE) مقدراً غير متحيز للتباين المشترك (σ^2) ، ولكن إذا كانت الفرضية خاطئة فإن E (MSTRT.) > E(MSE) ، وبالتالي سوف نرفض H_o إذا كانت : $F = \frac{MSTRT.}{MSE}$ أكبر من $f_{\alpha, t-1, N-t}$.

مثال (8): شركة منتجة للأدوية ترغب في تحديد مدى فعالية سائل دوائي إذا تم خلطه في أوعية كبيرة خصصت لتكريره ، وللقيام بهذه المهمة اختيرت عينة عشوائية من 5 أوعية من إنتاج عدة أشهر ومن كل وعاء تم اختيار 4 عينات عشوائية فكانت النتائج كما يلي :

الوعـــــاء

1	2	3	4	5
3.1	2.6	3.4	4.2	1.8
3.8	2.9	3.9	4.4	2.3
3.5	2.8	3.3	4.3	1.9
3.0	2.0	3.1	4.2	2.1
13.5	10.3	13.7	17.1	8.1

والمطلوب

أ - تكوين جدول تحليل لتباين واختبار الفرضية $H_0: \sigma_\alpha^2 = 0$ مقابل $H_1: \sigma_\alpha^2 > 0$ حيث σ_α^2 تمثل الاختلاف الناتج عن الاختلاف بين الأوعية عند مستوى المعنوية 5 ٪ .

ب - أختبر صحة فرضية تجانس التباين عند مستوى المعنوية 5 ٪ .

الحل :

أ - الحظ أن النموذج المناسب لهذه البيانات هو نموذج التأثيرات المتغيرة ، وذلك لأن الأوعية الخمسة تم اختيارها بطريقة عشوائية من بين مجموعة كبيرة من الأوعية ، وبالتالي إذا تم تكرار التجربة من الممكن استخدام أوعية أخرى .

الفرضية :

$$H_0: \sigma_\alpha^2 = 0$$
$$H_1: \sigma_\alpha^2 > 0$$

إحصاءة الاختبار :

إن الصيغ الرياضية لحساب إحصاءة الاختبار كما هي في حالة نموذج التأثيرات الثابتة ، وعليه فإن :

$$y_{..} = \sum_{i=1}^{5}\sum_{j=1}^{4} y_{ij} = 3.2 + 3.8 + \cdots 1.9 + 2.1 = 62.7 \Rightarrow \bar{y}_{..} = \frac{62.5}{20} = 3.135$$

$$\Rightarrow y_{..}^2 = (62.7)^2 = 3931.29 \quad \Rightarrow \quad \frac{y_{..}^2}{N} = \frac{3931.29}{20} = 196.5645$$

$$\frac{\sum_{i=1}^{5} y_{i.}^2}{5} = \frac{1}{5}[(13.5)^2 + (10.3)^2 + (13.7)^2 + (17.1)^2 + (8.1)^2] = 208.5125$$

$$SSTRT. = \frac{1}{n}\sum_{i=1}^{5} y_{i.}^2 - \frac{y_{..}^2}{tn} = 208.125 - 196.5645 = 11.94$$

$$SSE = \sum_{i=1}^{5}\sum_{j=1}^{4} y_{ij}^2 - \frac{1}{n}\sum_{i=1}^{5} y_{i.}^2 = 209.89 - 208.5125 = 1.3775$$

$$SST = \sum_{i=1}^{5}\sum_{j=1}^{4} y_{ij}^2 - \frac{y_{..}^2}{tn} = 209.89 - 196.5645 = 13.3255$$

وعليه فإن :

$$MSTRT = \frac{SSTRT}{t-1} = \frac{11.948}{4} = 2.981$$

$$MSE = \frac{SSE}{t(n-1)} = \frac{1.3775}{5 \times 3} = \frac{1.3775}{15} = 0.0918$$

وبالتالي يمكن وضع هذه النتائج في جدول تحليل التباين (ANOVA) على النحو الآتي :

المصدر S.V.	درجات حرية d.f.	مجموع المربعات SS	متوسط المربعات MS	F المحسوبة
المعالجات TRT	4	11.948	2.987	$\frac{2.987}{0.0918} = 32.5263$
الخطأ Error	15	1.3775	0.0918	
المجموع Total	19	13.3255		

القرار :

حيث أن $\alpha = 0.05$ ودرجات الحرية تساوى 4 و 15 وعليه من جدول F نجد أن $f_{0.05,4,15} = 3.06$

وحيث أن $32.5263 > 3.06$ وبالتالي نرفض H_0 عند المستوى 0.05 .

ب - لاختبار مدى صحة تجانس التباين بهذا المثال نستخدم اختبار بارتليت الذي سبق وأن أشرنا إليه وذلك كما يلي :

الفرضية :

$$H_0: \sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma_4^2 = \sigma_5^2$$

H_1: يختلف σ_i^2 على الأقل أحد

إحصاءة الاختبار :

حيث أن تباين كل معالجة يتم حسابه كما يلي : $s_i^2 = \frac{\sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{i.})^2}{n_i - 1}$ ، وعليه فإن :

المجتمع	$n_i - 1$	s_i^2	$(n_i - 1)\ln s_i^2$
1	3	0.1225	-6.2989326
2	3	0.162499	-5.4512337
3	3	0.115833	-6.4668123
4	3	0.009167	-14.076566
5	3	0.049167	-9.0376221
المجموع	15		-41.331216

حيث أن $s_p^2 = \frac{\sum_{i=1}^{t}(n_i - 1)s_i^2}{N-t} = \frac{SSE}{15} = MSE = 0.0918$ ، وعليه فإن :

وإن $\ln s_p^2 = -2.38778$

$$c = 1 + \frac{1}{3(t-1)}\left[\sum_{i=1}^{t}\left(\frac{1}{n_i - 1}\right) - \left(\frac{1}{N-t}\right)\right]$$

$$= 1 + \frac{1}{3(5-1)}\left[\left(\tfrac{1}{3}+\tfrac{1}{3}+\tfrac{1}{3}+\tfrac{1}{3}+\tfrac{1}{3}\right) - \tfrac{1}{15}\right] = 1.133333$$

$$w = (N-t)\ln s_p^2 - \sum_{i=1}^{t}(n_i - 1)\ln s_i^2$$

$$= (20-5)(-2.38778) - (-41.331216) = 5.514512$$

ومما سبق نجد أن :

$$B = \frac{w}{c} = \frac{5.514512}{1.133333} = 4.865746$$

القرار :

من جدول مربع كاى نجد أن $\chi^2_{\alpha,t-1} = \chi_{0.05,4} = 9.488$ ، وحيث أن $4.865746 < 9.488$ وعليه لا توجد معلومات كافية لرفض H_0 .أي يمكن القول بأن تباينات المجتمعات الخمسة متساوية .

مما سبق يتضح أن :

1 - في حالة نموذج التأثير الثابت إن المعالجات يتم اختيارها من قبل الباحث بينما في حالة نموذج التأثير المتغير يمكن التفكير فيها كعينة عشوائية من مجتمعاتها .

2 - في حالة نموذج التأثير الثابت إذا كرر الباحث التجربة فإنه سوف يستخدم نفس المعالجات بينما في حالة نموذج التأثير المتغير من الممكن استخدام معالجات مختلفة .

3 - حسابات مجموع المربعات بنفس الكيفية في النموذجيين .

4 - النموذجيين مختلفين والفرضيات مختلفة وتفسير النتائج مختلف .

10 - 5 التصميم العشوائي الكامل بقطاعات Randomized complete block design (RCBD)

لقد أشرنا في البند السابق إلى أن أحد الشروط الواجب توفرها في التصميم العشوائي الكامل هو أن تكون الوحدات التجريبية متجانسة ، وذلك لأنه إذا لم يتحقق هذا الشرط فإن ذلك سيؤدى إلى إخفاء الفروق الجوهرية بين المعالجات ، مما يؤدى إلى زيادة فرصة عدم رفض فرض العدم الذي ينص على وجود فروق بين المعالجات وفي الحقيقة هناك فروق ، وعليه إذا توفرت عند الباحث معلومات عن عدم تجانس الوحدات التجريبية فإنه يفضل أن يقسم هذه الوحدات إلى مجموعات متجانسة ثم تتم المقارنة بين المعالجات داخل هذه المجموعة ، ومن خلال هذا الأجراء سيكون التباين بين الوحدات التجريبية داخل كل مجموعة أقل من التباين الذي بين جميع الوحدات التجريبية .

إن التصميم العشوائي الكامل بقطاعات هو التصميم الذي يتم فيه تقسيم الوحدات التجريبية (Experimental units). التي سيتم تطبيق المعالجات عليها - إلى مجموعات متجانسة تسمى قطاعات (Blocks) ، بحيث يكون عدد الوحدات التجريبية بكل قطع يساوى عدد المعالجات المدروسة ، إن المعالجات يتم تصنيفها بطريقة عشوائية للوحدات التجريبية داخل كل قطاع ، ويجب أن يكون واضحاً هنا أن كل معالجة تظهر في كل قطاع وكل قطاع يعطى كل معالجة .إن كفاءة هذا التصميم تعتمد على قدرة الباحث في الحصول على قطاعات متجانسة للوحدات التجريبية وهذا بدوره معتمداً على مدى معرفة الباحث بالوحدات التجريبية ، وعند استخدام القطاعات بكفاءة سيؤدى ذلك إلى أن يكون متوسط مربع الخطأ صغير وهذا سيؤدى بدوره إلى زيادة فرصة رفض فرض العدم الذي ينص على عدم وجود فروق معنوية بين المعالجات ، ومن خصائصه أنه سهل الفهم وحساباته بسيطة . ولتوضيح كيفية تصميم تجربة بطريقة التصميم العشوائي الكامل بقطاعات ، نفترض أنه في المثال الذي تطرقنا إليه في بند (2 - 10) أن الحيوانات يمكن تقسيمها إلى أربعة مجموعات أكثر تجانساً حسب فئات العمر وإن فئات العمر هذه هي من 6 أشهر إلى أقل من سنة ومن سنة إلى أقل من سنتين ومن سنتين إلى أقل من 3

سنوات ومن 3 سنوات إلى أربعة سنوات حيث تتضمن كل فئة من فئات العمر ثلاثة حيوانات وبطريقة عشوائية سيتم توزيع الفيتامينات الثلاثة عشوائياً داخل كل فئة عمر (قطاع) بطريقة مستقلة عن فئات العمر (القطاعات) الأخرى .وبالرجوع إلى جدول الأرقام العشوائية بآخر الكتاب واختيار الأعمدة 11 - 14 وتسجيل الرقميين العشوائيين الأولين تم ترتيبها ترتيباً تصاعدياً وإعطاء كل معالجة للوحدة التجريبية حسب ترتيبها الخاص بها نتحصل على الآتي :

فئات العمر	الفيتامين	الرقم العشوائي	الترتيب	القطاع
1	A	22	1	A
	B	75	3	C
	C	44	2	B
2	A	20	2	B
	B	64	3	C
	C	16	1	A
3	A	94	3	C
	B	30	2	B
	C	14	1	A
4	A	88	2	B
	B	87	1	A
	C	92	3	C

وبالتالي فإن الشكل النهائي للتصميم كما يلي :

القطاع 4	القطاع 3	القطاع 2	القطاع 1
2 B	3 C	2 B	1 A
1 A	2 B	3 C	3 C
3 C	1 A	1 A	2 B

وبصفة عامة ستكون البيانات الناتجة من تجربة بتصميم عشوائي كامل بقطاعات كما هي موضحة بالجدول التالي :

المعالجات	القطاعات 1	2	$\cdots$	j	$\cdots$	b	المجموع	المتوسط
1	y_{11}	y_{12}	$\cdots$	y_{1j}	$\cdots$	y_{1b}	$y_{1.}$	$\bar{y}_{1.}$
2	y_{21}	y_{22}	$\cdots$	y_{2j}	$\cdots$	y_{2b}	$y_{2.}$	$\bar{y}_{2.}$
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$	$\vdots$
i	y_{i1}	y_{i2}	$\cdots$	y_{ij}	$\cdots$	y_{ib}	$y_{i.}$	$\bar{y}_{i.}$
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$	$\vdots$
t	y_{t1}	y_{t2}	$\cdots$	y_{tj}	$\cdots$	y_{tb}	$y_{t.}$	$\bar{y}_{t.}$
المجموع	$y_{.1}$	$y_{.2}$	$\cdots$	$y_{.j}$	$\cdots$	$y_{.b}$	$y_{..}$	
المتوسط	$\bar{y}_{.1}$	$\bar{y}_{.2}$	$\cdots$	$\bar{y}_{.j}$	$\cdots$	$\bar{y}_{.b}$		$\bar{y}_{..}$

حيث :

مجموع جميع المشاهدات أو القراءات أو القياسات = $y_{..} = \sum_{i=1}^{t}\sum_{j=1}^{b} y_{ij}$

مجموع المشاهدات بالمعالجة i = $y_{i.} = \sum_{j=1}^{b} y_{ij} \quad \Leftarrow \quad \bar{y}_{i.} = \frac{y_{i.}}{b}$ = متوسط المعالجة i .

مجموع المشاهدات بالقطاع j = $y_{.j} = \sum_{i=1}^{t} y_{ij} \quad \Leftarrow \quad \bar{y}_{.j} = \frac{y_{.j}}{t}$ = متوسط القطاع j .

إن الأسلوب المتبع في التحليل البيانات الناتجة من التصميم العشوائي الكامل بقطاعات يطلق عليه تسمية تحليل التباين الثنائي (Two - way analysis of variance) ، وذلك لأن المشاهدات تم تصنيفها بناءً على معيارين هما القطاع الذي تقع فيه ومجموعة المعالجة التي تنتمي إليها .إن نموذج التصميم العشوائي الكامل بقطاعات يكون على النحو التالي :

$$y_{ij} = \mu + \alpha_i + \beta_j + \epsilon_{ij} \quad , \quad \begin{cases} i = 1, 2, \ldots, t \\ j = 1, 2, \ldots, b \end{cases} \qquad (33)$$

حيث : y_{ij} المفردة أو المشاهدة أو القراءة بالقطاع j الناتجة من تطبيق المعالجة i عليها .

و μ ترمز للمتوسط العام و α_i تمثل تأثير المعالجة i و β_j تمثل تأثير القطاع j و ϵ_{ij} الخطأ العشوائي الذي يمثل جميع مصادر الاختلاف غير الناتجة عن القطاعات أو المعالجات .

ويقوم هذا النموذج على الافتراضات التالية :

1 - الأخطاء العشوائية ϵ_{ij} مستقلة عن بعضها البعض ولها توزيع طبيعي بمتوسط يساوى صفر وتباين يساوى σ^2 أى أن $\epsilon_{ij} \sim NID(0, \sigma^2)$.

2 - إن تأثير كل من القطاع والمعالجة تجميعي بمعنى أن أى توفيق لمعالجة وقطاع لا ينتج تأثير أكبر من أو أقل من مجموع تأثير كل منها على حده ، وفي هذه الحالة يمكن الإثبات بأن $\sum_{i=1}^{t} \alpha_i = \sum_{j=1}^{b} \beta_j = 0$ أى ان تأثير كل من القطاع والمعالجة ثابت ، وبعبارة أخرى إن القيم المتوقعة للمفردات في قطاعات مختلفة لنفس المعالجة من الممكن أن تختلف ولكن تأثير المعالجات متساوي بجميع القطاعات .

3 - إذا كان تأثير المعالجات عشوائي فإن التغيير الوحيد في النموذج أعلاه هو أن $\alpha_i \sim NID(0, \sigma_\alpha^2)$ و α_i مستقلة عن ϵ_{ij} . وبالمثل إذا كان تأثير القطاعات عشوائي فإن التغير الوحيد في النموذج أعلاه هو أن $\beta_i \sim NID(0, \sigma_\beta^2)$ و β_j مستقلة عن ϵ_{ij} ،وإذا صح هذا الافتراض الأخير أي أن القطاعات عشوائية فإنه ليس بالضرورة استخدام النموذج الذي يفترض أنه لا يوجد تفاعل ما بين القطاعات والمعالجات حيث من الممكن استخدام النموذج الذي من خلاله يمكن الاختبار للتفاعل فيما بينهما وفي أي من الحالتين التي يكون فيهما إحداهما ثابت والأخر عشوائي فإن النموذج يطلق عليه تسمية نموذج مختلط (Mixed effects model) .

إذا كان تأثير كل من المعالجات والقطاعات ثابت فإنه يمكن تقدير المعلمات في التصميم العشوائي الكامل بقطاعات باستخدام طريقة المربعات الصغرى بتطبيق القيدين $\sum_{i=1}^{t} \alpha_i = 0$ و $\sum_{j=1}^{b} \beta_j = 0$ وإن التقدير بقيمة واحدة لهذه المعلمات وفقاً لطريقة المربعات الصغرى سيكون كما يلي :

$$
\begin{aligned}
\hat{\mu} &= \bar{y}_{..} \\
\hat{\alpha}_i &= \bar{y}_{i.} - \bar{y}_{..} \quad , \quad i = 1, 2, \cdots, t \\
\hat{\beta}_j &= \bar{y}_{.j} - \bar{y}_{..} \quad , \quad j = 1, 2, \cdots, b
\end{aligned}
\qquad (34)
$$

إن أحد الأهداف الرئيسية من وراء استخدام التصميم العشوائي الكامل بقطاعات هو اكتشاف فيما إذا كان هناك فروق معنوية ما بين متوسطات المعالجات ، وعليه إذا كان :

1 - تأثير كل من المعالجات والقطاعات ثابت فإن الفرضية تكون كالآتي :

$H_o: \alpha_1 = \alpha_2 = ... = \alpha_t = 0$

$H_1: (\alpha_i \neq 0)$ على الأقل واحدة لا تساوى 0

2 - تأثير المعالجات عشوائي فإن الفرضية تكون كالأتي :

$$H_o: \sigma_\alpha^2 = 0$$
$$H_1: \sigma_\alpha^2 > 0$$

3 - تأثير القطاعات عشوائي فإن الفرضية ستكون كالأتي :

$$H_o: \sigma_\beta^2 = 0$$
$$H_1: \sigma_\beta^2 > 0$$

أما في ما يخص تأثير القطاعات أي $H_o: \beta_1 = \beta_2 = ... = \beta_b = 0$ غالباً ما تستخدم لتحديد ما إذا كان من الضروري استخدام القطاعات في تجارب مستقبلية مشابهة ، ويجب أن يكون الباحث حذر جداً في اختبار هذه الفرضية وهناك جدال كبير حول هذه الفرضية ما بين الإحصائيين ويوصى معظمهم بعدم إجراء مثل هذا الاختبار وسوف لن نتعرض لمثل هذا الاختبار هنا .

إن الاختبار الذي سوف يستخدم لصحة أي من الفرضيات أعلاه سيكون مبنى على مقارنة مقدرين مستقلين لتباين المجتمع المشترك (σ^2) ، ويمكن الحصول على هذين المقدرين من خلال تجزئة الاختلاف الكلى في البيانات وذلك كما يلي :

$$\sum_{i=1}^{t}\sum_{j=1}^{n}(y_{ij} - \bar{y}_{..})^2 = \sum_{i=1}^{t}\sum_{j=1}^{n}[(y_{i.} - \bar{y}_{..}) + (\bar{y}_{.j} - \bar{y}_{..}) + (y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} - \bar{y}_{..})]^2$$

وبتبسيط الطرف الأيمن من هذه المعادلة نجد أن :

$$\sum_{i=1}^{t}\sum_{j=1}^{n}(y_{ij} - \bar{y}_{..})^2 = b\sum_{i=1}^{t}(y_{i.} - \bar{y}_{..})^2 + t\sum_{j=1}^{b}(\bar{y}_{.j} - \bar{y}_{..})^2$$
$$+ \sum_{i=1}^{t}\sum_{j=1}^{b}(y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} - \bar{y}_{..})^2 + 2\sum_{i=1}^{t}\sum_{j=1}^{b}(y_{i.} - \bar{y}_{..}).(\bar{y}_{.j} - \bar{y}_{..})$$
$$+ 2\sum_{i=1}^{t}\sum_{j=1}^{b}(y_{i.} - \bar{y}_{..})(y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} - \bar{y}_{..})$$
$$+ 2\sum_{i=1}^{t}\sum_{j=1}^{b}(\bar{y}_{.j} - \bar{y}_{..})(y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} - \bar{y}_{..})$$

وحيث أن التقاطعات الضربية التي بالطرف الأيمن جميعها تساوى صفر ، وبالتالي نجد أن :

$$\sum_{i=1}^{t}\sum_{j=1}^{b}(y_{ij}-\bar{y}_{..})^2 = b\sum_{i=1}^{t}(y_{i.}-\bar{y}_{..})^2 + t\sum_{j=1}^{b}(\bar{y}_{.j}-\bar{y}_{..})^2 + \sum_{i=1}^{t}\sum_{j=1}^{b}(y_{ij}-\bar{y}_{i.}-\bar{y}_{.j}-\bar{y}_{..})^2$$

أي أن :

$$SST = SSTRT. + SSB + SSE$$

وحيث أنه هناك $N = bt$ مفردة وبالتالي فإن درجات الحرية المصاحبة لمجموع المربعات الكلية تساوى $N - 1$ ودرجات الحرية المصاحبة إلى .SSTRT تساوى $t - 1$ ، ودرجات الحرية المصاحبة إلى SSB تساوى $b - 1$ ودرجات الحرية المصاحبة إلى SSE تساوى $bt - 1 - (t - 1) - (b - 1) = (t - 1)(b - 1)$ ، وبقسمة كل مجموع مربعات على درجات الحرية المصاحبة له نتحصل على متوسط المربعات ، أي أن :

$$MSE = \frac{SSE}{(t-1)(b-1)} \quad ، \quad MSTRT. = \frac{SSTRT.}{t-1} \quad ، \quad MSB = \frac{SSB}{b-1}$$

إذا كان تأثير المعالجات والقطاعات ثابت فإن القيمة المتوقعة لمتوسط المربعات تكون كما يلي :

$$E(MSTRT.) = \sigma^2 + \frac{b}{t-1}\sum_{i=1}^{t}\alpha_i^2$$

$$E(MSB) = \sigma^2 + \frac{t}{b-1}\sum_{j=1}^{b}\beta_j^2$$

$$E(MSE) = \sigma^2$$

وإذا كانت الفرضية (1) صحيحة فإن $E(MSTRT.) = E(MSE)$ وخلاف ذلك يكون $E(MSTRT.)$ أكبر من $E(MSE)$.أما إذا كان تأثير المعالجات ثابت والقطاعات عشوائي فإن :

$$E(MSTRT.) = \sigma^2 + \frac{b}{t-1}\sum_{i=1}^{t}\alpha_i^2$$

$$E(MSB) = \sigma^2 + t\sigma_\beta^2$$

$$E(MSE) = \sigma^2$$

وبالمثل إذا كانت الفرضية (2) صحيحة فإن $E(MSTRT.) = E(MSE)$ وخلاف ذلك يكون $E(MSTRT.)$ أكبر من $E(MSE)$.

وأخيراً إذا كان تأثير المعالجات عشوائي والقطاعات عشوائي فإن :

$$E(MSTRT.) = \sigma^2 + b\sigma_\alpha^2$$
$$E(MSB) = \sigma^2 + t\sigma_\beta^2$$
$$E(MSE) = \sigma^2$$

وإذا كانت الفرضية (3) صحيحة فإن $E(MSB) = E(MSE)$ ، وخلاف ذلك يكون $E(MSB)$ أكبر من $E(MSE)$. وعليه لاختبار الفرضيـة (1) أو (2) سـوف نرفض H_o إذا كانت :

$$F_T = \frac{MSTRT.}{MSE} > f_{\alpha\ ,t-1,(t-1)(b-1)}$$

ولاختبار الفرضية (3) سوف نرفض H_o إذا كانت :

$$F_B = \frac{MSB.}{MSE} > f_{\alpha\ ,b-1,(t-1)(b-1)}$$

وكما سبق عادةً ما يتم وضع النقاط السابقة في جدول يطلق عليه تسمية جدول تحليل التباين وذلك كما يلي :

جدول تحليل التباين الثنائي (2- way ANOVA)

مصدر الاختلاف S. V.	مجموع المربعات SS	درجات الحرية df	متوسط الخطأ MS	F المحسوبة
TRT. :المعالجات	SSTRT.	t - 1	$MSTRT. = \frac{SSTRT.}{t-1}$	$F_T = \frac{MSTRT.}{MSE}$
Blocks :القطاعات	SSB	b - 1	$MSB = \frac{SSB}{(b-1)}$	$F_B = \frac{MSB}{MSE}$
Error : الخطأ	SSE	(t - 1) (b - 1)	$MSE = \frac{SSE}{(t-1)(b-1)}$	
Total المجموع	SST	t b - 1		

وإنه من المناسب من الناحية الحسابية إعادة كتابة صيغ مجاميع المربعات وذلك كما يلي :

$$SST = \sum_{i=1}^{t}\sum_{j=1}^{b}(y_{ij} - \bar{y}_{..})^2 = \sum_{i=1}^{t}\sum_{j=1}^{b} y_{ij}^2 - \frac{y_{..}^2}{N} \quad (35)$$

$$SSTRT. = \sum_{i=1}^{t}\sum_{j=1}^{n}(\bar{y}_{i.} - \bar{y}_{..})^2 = \frac{1}{b}\sum_{i=1}^{t} y_{i.}^2 - \frac{y_{..}^2}{N} \quad (36)$$

$$SSB = \sum_{i=1}^{t}\sum_{j=1}^{b}(\bar{y}_{.j} - \bar{y}_{..})^2 = \frac{1}{t}\sum_{j=1}^{b} y_{.j}^2 - \frac{y_{..}^2}{N} \quad (37)$$

$$SSE = \sum_{i=1}^{t}\sum_{j=1}^{b}(y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} - \bar{y}_{..})^2 \quad (38)$$

$$= SST - SSTRT. - SSB$$

لقد أشرنا في تحليل التباين الأحادي عند رفض فرض العدم يمكن إجراء مقارنات متعددة لتحديد اى المتوسطات تختلف عن بعضها البعض إن هذا الأسلوب يمكن عمله أيضاً في حالة التصميم العشوائي الكامل بقطاعات .

مثال (9) : أجريت تجربة لمقارنة تأثير 3 أنواع من المبيدات الحشرية على نوع معين من البذور حيث تم اختيار 4 قطاعات وقسم كل منها إلى 3 صفوف وتركت مسافة مناسبة بين الصفوف الثلاثة داخل كل قطاع وبكل قطاع تم زرع 100 بذرة وتم حفظها تحت المبيد الذي خصص للصف ، وبطريقة عشوائية صنفت المبيدات للصفوف بداخل القطاع بحيث كل مبيد ظهر في كل صف في القطاعات الأربعة والمستجيب الذي يهم الباحث هنا هو عدد النباتات التي تنشأ بكل صف ، فكانت النتائج كما يلي :

	القطاعات			
المبيد	1	2	3	4
I	56	49	65	60
II	84	87	94	93
III	80	72	83	85

والمطلوب اختبار الفرضيات التالية عند مستوى المعنوية 1 ٪ :

أ - $H_0 : \alpha_1 = \alpha_2 = \alpha_3 = 0$

$H_1 : \alpha_i \neq 0$ على الأقل واحدة لا تساوى

ب -

$$H_0 : \sigma^2_\beta = 0$$
$$H_1 : \sigma^2_\beta > 0$$

الحل :

من الجدول اعلاه نجد أن :

$y_{1.} = 230 \quad y_{2.} = 349 \quad y_{3.} = 320$

$y_{.1} = 220 \quad y_{.2} = 199 \quad y_{.3} = 242 \quad y_{.4} = 238 \quad y_{..} = 899$

$$\frac{\sum_{i=1}^{4} y_{i.}^2}{4} = 69275.25 \text{ ، } \frac{\sum_{j=1}^{3} y_{.j}^2}{3} = 67736.333 \text{ , } \frac{y_{..}^2}{12} = 67350.083$$

$$\sum_{i=1}^{3}\sum_{j=1}^{4} y_{ij}^2 = 69685$$

وعليه فإن :

$$SST = \sum_{i=1}^{t}\sum_{j=1}^{b} y_{ij}^2 - \frac{y_{..}^2}{N} = 69685 - 67350.083 = 2334.917$$

$$SSTRT. = \frac{1}{b}\sum_{i=1}^{t} y_{i.}^2 - \frac{y_{..}^2}{N} = 69275.25 - 67350.083 = 1925.167$$

$$SSB = \frac{1}{t}\sum_{j=1}^{b} y_{.j}^2 - \frac{y_{..}^2}{N} = 67736.333 - 67350.083 = 386.25$$

$$SSE = SST - SSTRT. - SSB$$
$$= 2334.917 - 1925.167 - 386.250 = 23.500$$

وعليه يمكن وضع النتائج السابقة في جدول تحليل التباين الثنائي التالي :

مصدر الاختلاف S.V.	df	SS	MS	F المحسوبة
TRT. المبيد	2	1925.167	962.584	$F_T = \frac{MSTRT.}{MSE} = 245.56$
Block القطاع	3	386.250	128.750	$F_B = \frac{MSB}{MSE} = 32.86$
Error الخطأ	6	23.500	3.917	
Total الكلي	11	2334.917		

وبالتالي فإن :

أ - من جدول F وبدرجات حرية تساوى 2 و 6 و $\alpha = 0.01$ نجد أن $f_{0.01,2,6} = 10.92$

وحيث أن :

$F_T = 245.56 > 10.92$ وعليه نرفض H_o ، أي أن للمبيدات تأثير مختلف .

ب - من جدول F وبدرجات حرية تساوى 3 و 6 و $\alpha = 0.01$ نجد أن $f_{0.01,3,6} = 9.78$

وحيث أن :

$F_B = 32.86 > 9.78$ وعليه نرفض H_o .

إن أسلوب المقارنات المتعددة الذي تعرضنا إليه في حالة تحليل التباين الأحادي يمكن عمله أيضاً في حالة تحليل التباين الثنائي ، وعليه إذا كانت L تمثل علاقة خطية في متوسطات المعالجات فإنه يمكن اختبار الفرضية الآتية :

$$H_0 : L = 0$$
$$H_1 : L \neq 0$$

وذلك باستخدام اختبار t المألوف على النحو الآتي :

$$t = \frac{\hat{L}}{\sqrt{(MSE)\sum_{i=1}^{t}\frac{c_i^2}{b}}} \qquad (39)$$

حيث $\hat{L} = \sum_{i=1}^{t} c_i\,\bar{y}_{i.}$ و MSE يمثل متوسط مربع الخطأ بجدول تحليل التباين ، وسوف نرفض H_0 عند مستوى المعنوية α إذا كانت $|t| \geq t_{\frac{\alpha}{2},(b-1)(t-1)}$.

أما في ما يخص اختبار توكى (HSD) واختبار فيشر (LSD) فيكونا كما هما بدون تغيير عدا تعويض b بدلاً من n في الصيغتين .

مثال (10) : من بيانات المثال السابق : أ - أختبر في ما إذا كان متوسط المبيدين الثاني والثالث يختلف عن متوسط المبيد الأول . ب - أستخدم اختبار توكى لاكتشاف أى المبيدات تختلف عن بعضها البعض .

الحل :

من البيانات السابقة نجد أن $\bar{y}_1 = 57.5$ ، $\bar{y}_2 = 87.25$ ، $\bar{y}_3 = 80.0$

أ - إن الفرضية المطلوب اختبارها تكون على النحو الآتي :

$$H_0 : L = 0$$
$$H_1 : L \neq 0$$

حيث $L = \mu_1 - \left(\frac{\mu_2 + \mu_3}{2}\right) = \sum_{i=1}^{3} c_i \mu_i$ و $c_1 = 1 ، c_2 = c_3 = -\frac{1}{2}$

وعليه فإن :

$$\hat{L} = \sum_{i=1}^{3} c_i \hat{\mu}_i = \sum_{i=1}^{3} c_i \bar{y}_i = 57.5 - \left(\frac{87.25 + 80.0}{2}\right) = -26.125$$

وإن

$$\sqrt{(MSE)\sum_{i=1}^{t} \frac{c_i^2}{b}} = \sqrt{(3.917)\{\tfrac{1}{4}[(-1)^2 + (-\tfrac{1}{2})^2 + (-\tfrac{1}{2})^2]\}} = 1.212$$

وبالتالي فإن :

$$t = \frac{\hat{L}}{\sqrt{(MSE)\sum_{i=1}^{t} \frac{c_i^2}{b}}} = \frac{-26.125}{1.212} = -21.555$$

ومن جدول t وبدرجات حرية تساوي 6 و $\frac{\alpha}{2} = 0.025$ نجد أن $t_{0.025,6} = 2.441$ وحيث أن $2.441 < |t|$ وعليه نرفض H_0 .

ب - حيث أنه من جدول (13) نجد أن $q_{\alpha,t,(b-1)(t-1)} = q_{0.05,3,6} = 4.34$ وإن

$$HSD = q_{\alpha,t,(b-1)(t-1)}\sqrt{\frac{MSE}{b}} = (4.34)\sqrt{\frac{3.917}{4}} = 4.295$$

وبالتالي فإن :

الفرضية	الفرق المطلق ما بين المتوسطين	HSD	القرار
$H_o:\mu_1 = \mu_2$	29.75	4.295	نرفض H_o
$H_o:\mu_1 = \mu_3$	22.5	4.295	نرفض H_o
$H_o:\mu_2 = \mu_3$	2.75	4.295	لا نرفض H_o

إن نتائج الجدول السابق يمكن تلخيصها بالطريقة المقترحة من قبل ونكن (Duncam) وذلك بوضع خط يربط ما بين أى متوسطين أو أكثر لا يختلفان .

$$\mu_1 \quad \underline{\mu_3 \quad \mu_2}$$

وعليه يمكن القول بأن μ_2 و μ_1 يختلفان و μ_3 و μ_1 يختلفان أيضاً ، ولكن μ_3 و μ_2 متساويان ·

10 - 6 التجارب العامليه Factorial Experiments

إن تصميم التجارب الذي سبق وان تعرضنا إليها تهتم بتأثير متغير واحد فقط ، ولكن في بعض الأحيان قد يكون من المرغوب فيه دراسة تأثير متغيرين أو أكثر في آن واحد ، فمثلاً دراسة تأثير مستويات مبيد حشري معين ومستويات سماد معين على محصول زراعي معين ، أو دراسة تأثير الحرارة والرطوبة على معدل نمو النبات ، أو دراسة تأثير طرائق مختلفة للتدريس ومستويات مختلفة من الطلبة على مستوى التحصيل العلمي ، ... الخ .إن التجارب التى يتم فيها تحليل تأثير متغيرين أو أكثر في آن واحد تسمى تجارب عامليه ، ويمكن القول بأن هذا النوع من التجارب هو أكفاء التصميمات لمثل هذا النوع من التحليل .

إن التجربة العامليه هي كل تجربة تكون فيها المعالجات عبارة عن مجموعة من التوافيق بين عدة مستويات (levels) لعدة عوامل (Factors) اى أن مستوى كل عامل في التجربة يظهر مع مستويات كل العوامل الأخرى ، ويعرف العامل على انه نوع من المعالجة التى تحتوى على تقسيمات متعددة تسمى بالمستويات ، مثل استعمال أربعة جرعات مختلفة من حيث التركيز

أو مثلاً عامل لدواء معين ، أو مثلاً ثلاثة مستويات مختلفة لدرجات الحرارة 20 ، 30 ، 40 ، أو مثلاً عامل التربة في الأبحاث الزراعية طينية ورملية وطمئية . وعادة ما يكون الاهتمام في التجارب العاملية بالتأثيرات الرئيسية للعوامل بالإضافة للتفاعل بينها ، حيث يعرف التأثير الرئيسي للعامل (main effects of a factor) على أنه التغير في المستجيب نتيجة لتغير مستوى العامل والسبب في تسميته بهذا الاسم وذلك لأنه يحظى بأكثر الاهتمام في التجربة ، أما التأثير البسيط للعامل (simple effect of a factor) فهو الفرق في الاستجابة بين مستويي عامل معين عند مستوى معين لعامل أخر ، وبالتالي فإن التأثير الرئيسي لعامل معين يساوى متوسط التأثيرات البسيطة له . أما التفاعل (Interaction) فهو الاختلاف في المستجيب بين مستويات عامل معين نتيجة لتغير مستويات عامل أخر .

ولتوضيح المفاهيم السابقة نفترض بأن البيانات التالية تمثل متوسطات الإنتاج لثلاثة أنواع من الحبوب (B) تم الحصول عليها باستخدام نوعين من الأسمدة (A) .

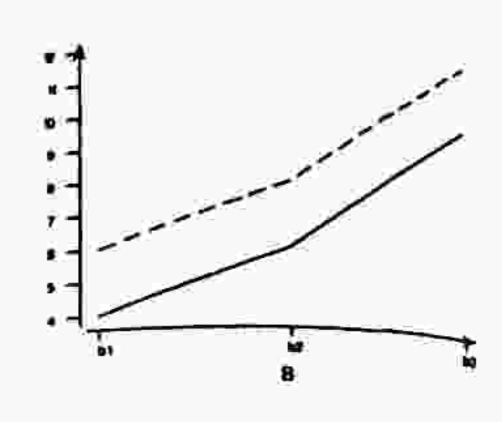
العامل A (الأسمدة)	العامل B (الحبوب) J = 1	J = 2	J = 3
السماد الأول I = 1	$\mu_{11} = 4$	$\mu_{12} = 6$	$\mu_{13} = 10$
السماد الثاني I = 2	$\mu_{21} = 6$	$\mu_{22} = 8$	$\mu_{23} = 12$

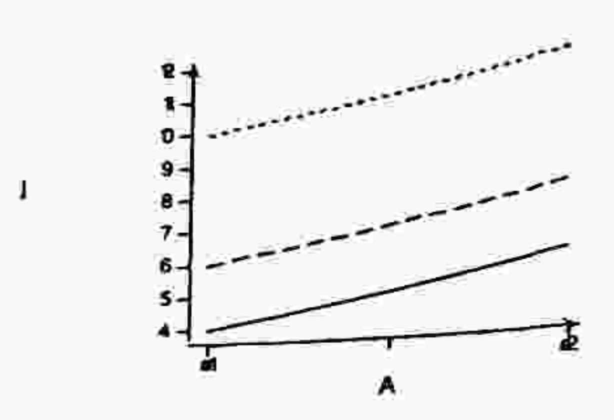
ومن خلال هذا الجدول يتضح الآتي :

أ - بالنسبة لمستويي العامل A إن الفرق ما بين متوسطي أى مستويين من مستويات العامل B متساوي ، أى أنه بالنسبة للمستوى الأول من العامل A يكون $\mu_{12} - \mu_{11} = 2$, $\mu_{13} - \mu_{12} = 4$, $\mu_{13} - \mu_{11} = 6$ وهى نفس الفروق بالنسبة للمستوى الثاني .

ب - بالنسبة لجميع مستويات العامل B يكون الفرق ما بين متوسطي أى مستويين من مستويات العامل A متساوي أى أن $\mu_{21} - \mu_{11} = \mu_{22} - \mu_{12} = \mu_{23} - \mu_{13} = 2$.

جـ - عند تمثيل هذه البيانات بيانياً كما في الشكل(1) نلحظ أن الخطوط متوازية لجميع مستويات كل عامل .





شكل (1) : عدم وجود تفاعل بين الأسمدة والحبوب .

عندما تتوفر في أي بيانات لتجربة من عاملين الخصائص الثلاثة السابقة يقال بأنه لا يوجد تفاعل ما بين العاملين .

إن وجود التفاعل من الممكن أن يؤثر على خصائص البيانات بطرائق عديدة وذلك اعتماداً على طبيعة التفاعل . وسوف نوضح أحد أنواع التفاعل من خلال إجراء بعض التغيير في بيانات الجدول السابق وذلك كما يلي :

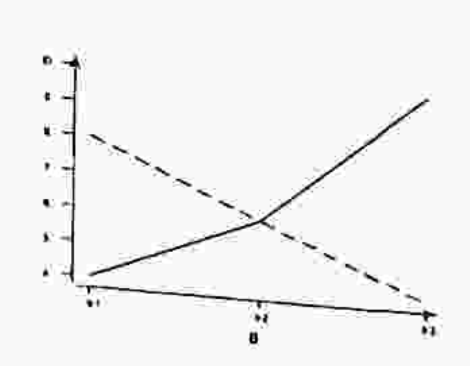
العامل A (الأسمدة)	العامل B (الحبوب)		
	$J = 1$	$J = 2$	$J = 3$
السماد الأول $I = 1$	$\mu_{11} = 4$	$\mu_{12} = 6$	$\mu_{13} = 10$
السماد الثاني $I = 2$	$\mu_{21} = 8$	$\mu_{22} = 6$	$\mu_{23} = 4$

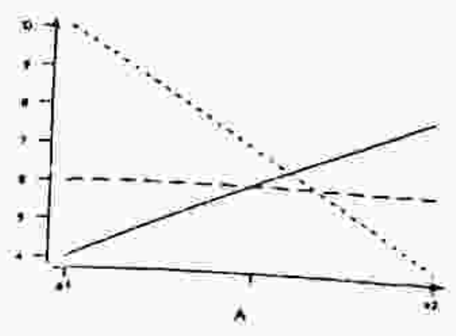
ومن خلال هذا الجدول يتضح أن :

أ - إن الفرق ما بين متوسطي أي مستويين من مستويات العامل B مختلف بالنسبة لمستويي العامل A .

ب - إن الفرق ما بين متوسطي مستويي العامل A غير متساوي بالنسبة لجميع مستويات العامل B .

جـ - من خلال رسم بيانات هذا الجدول كما في شكل (2) يتضح أن منحنيات مستوى أي عامل غير متوازية ، إن أي بيانات لتجربة من عاملين تتصف بهذه الخواص يقال بأنه يوجد تفاعل ما بين العاملين ، وعليه إذا وجد تفاعل يجب دراسة متوسطات التفاعل ما بين العاملين بدلاً من دراسة متوسط كل عامل على حده .





شكل (2) : وجود تفاعل بين الأسمدة والحبوب .

مما سبق يمكن القول بأن من مزايا التجارب العامليه ما يلي :

أ - يمكن دراسة التفاعلات وتقديرها .

ب - حيث أنه تتم دراسة عدة عوامل في تجربة واحدة ، وبالتالي فإن النتائج المتحصل عليها سيكون لها تطبيقات عديدة .

جـ - هناك ادخار في التكلفة والوقت وذلك لأنه عندما لا يوجد تفاعل بين العوامل فإن تأثيراتها سيتم تقديرها بدرجة عالية من الدقة .

إن التجارب العامليه يمكن دراستها باستخدام التصميم العشوائي الكامل أو التصميم العشوائي الكامل بقطاعات إلا أننا سوف نقتصر في دراستنا هنا على تحليل التجارب العاملية من وجهة التصميم العشوائي الكامل بعاملين ، فإذا رمزنا للعامل الأول بالرمز A وله " a " مستوى وللعامل الثاني بالرمز B وله " b " مستوى ، واعتبرنا كل توفيق (خلية) من مستويات العامل A مع مستويات العامل B على أنه معالجة ، وعليه سيكون هناك n مشاهدة بكل معالجة ، فإذا كانت y_{ijk} ترمز للمشاهدة k عند المستوى i للعامل A والمستوى j للعامل B حيث $i = 1,2,...,a$ و $j = 1,2,...,b$ و $k = 1,2,...,n$ ، فسيكون هناك ab من الخلايا وكل خلية تحتوى على n من المشاهدات ، وذلك كما هو موضح في الجدول الآتي :

A	B 1 2 ... b	المجموع	المتوسط
1	y_{111} y_{121} ⋯ y_{1b1} y_{112} y_{122} ⋯ y_{1b2} ⋮ ⋮ ⋮ y_{11n} y_{12n} ⋯ y_{2bn}	$y_{1..}$	$\bar{y}_{1..}$
2	y_{211} y_{221} ⋯ y_{2b1} y_{212} y_{222} ⋯ y_{2b2} ⋮ ⋮ ⋮ y_{21n} y_{22n} ⋯ y_{2bn} ⋮ ⋮ ⋮	$y_{2..}$	$\bar{y}_{2..}$
a	y_{a11} y_{a21} ⋯ y_{ab1} y_{a12} y_{a22} ⋯ y_{ab2} ⋮ ⋮ ⋮ y_{a1n} y_{a2n} ⋯ y_{abn}	$y_{a..}$	$\bar{y}_{a..}$
المجموع	$y_{.1.}$ $y_{.2.}$ ⋯ $y_{.b.}$	$y_{...}$	
المتوسط	$\bar{y}_{.1.}$ $\bar{y}_{.2.}$ ⋯ $\bar{y}_{.b.}$		$\bar{y}_{...}$

لحظ التشابه ما بين هذا الجدول والجدول الذي يعرض بيانات من تصميم عشوائي كامل بقطاعات ، ولكن لكي نختبر وجود التفاعل ما بين العاملين في التجارب العاملية يجب أن يكون هناك على الأقل مفردتين بكل خلية ، ولكن التصميم العشوائي الكامل بقطاعات يتطلب مفردة واحدة فقط . وسوف نستخدم تحليل التباين الثنائي هنا لتحليل البيانات الناتجة من التجارب بعاملين، وللوصول إلى ذلك سوف نعرف بالرموز الآتية :

$y_{ij.} = \sum_{k=1}^{n} y_{ijk}$ = مجموع المشاهدات التي في المستوى i للعامل A والمستوى j للعامل B

ومتوسطها $\bar{y}_{ij.} = \frac{y_{ij.}}{n}$.

$y_{i..} = \sum_{j=1}^{b}\sum_{k=1}^{n} y_{ijk}$ = مجموع مشاهدات المستوى i من العامل A ومتوسطها $\bar{y}_{i..} = \frac{y_{i..}}{bn}$.

$y_{.j.} = \sum_{i=1}^{a}\sum_{k=1}^{n} y_{ijk}$ = مجموع مشاهدات المستوى j من العامل B ومتوسطها $\bar{y}_{.j.} = \dfrac{y_{.j.}}{an}$.

$y_{...} = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n} y_{ijk}$ = مجموع جميع المشاهدات abn ومتوسطها $\bar{y}_{...} = \dfrac{y_{...}}{abn}$.

وسوف نفترض بأن المشاهدات في الخلية (i , j) تشكل عينة عشوائية حجمها n من مجتمع له توزيع طبيعي متوسطه μ_{ij} وتباينه σ^2 ، وإن جميع المجتمعات " ab " نفترض أن لها نفس التباين وهو σ^2 . إن كل مشاهدة في الجدول أعلاه يمكن كتابتها على النحو الآتي :

$$y_{ijk} = \mu_{ij} + \epsilon_{ijk} \qquad (40)$$

حيث " ϵ_{ijk} " يقيس انحرافات القيم y_{ijk} المشاهدة في الخلية (i , j) عن متوسط المجتمع μ_{ij} ، وإن $\epsilon_{ijk} \sim NID(0, \sigma^2)$ ، فإذا رمزنا لتأثير المستوى i من العامل A بالرمز α_i ولتأثير المستوى j من العامل B بالرمز β_j ولمتوسط التأثير الكلي بالرمز μ ولتأثير التفاعل بين المستوى i من العامل A والمستوى j من العامل B بالرمز $(\alpha\beta)_{ij}$. فإنه يمكننا كتابة متوسط المجتمع بالصورة الآتية :

$$\mu_{ij} = \mu + \alpha_i + \beta_i + (\alpha\beta)_{ij} \qquad (41)$$

وبالتالي فإن :

$$y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ijk} \quad , \begin{cases} i = 1,2,\ldots,a \\ j = 1,2,\ldots,b \\ k = 1,2,\ldots,n \end{cases} \qquad (42)$$

10 - 6 - 1 نموذج التأثيرات الثابتة Fixed Effects Model

وفقاً لهذا النموذج نفترض بأن العاملين A و B ثابتين ، اى أن مستويات العاملين وهما " a " و" b " من المستويات على التوالي قد تم اختيارها من قبل الباحث بشكل محدد ، وإن أى استنتاج يتوصل إليه الباحث سينطبق على هذه المستويات فقط ، وفي مثل هذا النوع من النماذج جرت العادة على تعريف التأثيرات α_i و β_j و $(\alpha\beta)_{ij}$ كانحرافات عن متوسطها ، وبذلك يكون $\sum_{i=1}^{a}\alpha_i = \sum_{j=1}^{b}\beta_j = \sum_{i=1}^{a}(\alpha\beta)_{ij} = \sum_{j=1}^{b}(\alpha\beta)_{ij} = 0$ ، ويمكن تقدير معلمات نموذج تحليل التباين باستخدام طريقة المربعات الصغرى وبتطبيق القيود :

$$\sum_{i=1}^{a}\alpha_i = 0\ ,\ \sum_{j=1}^{b}\beta_j = 0\ ,\ \sum_{j=1}^{b}(\alpha\beta)_{ij} = 0\ ,\ i = 1,2,\ldots,a\ \ ,\ \sum_{i=1}^{a}(\alpha\beta)_{ij} = 0\ ,\ i = 1,2,\ldots,b$$

فإن التقدير بقيمة واحدة لكل من μ و α_i و β_j و $(\alpha\beta)_{ij}$ ستكون كالآتي :

$$\begin{aligned}
\hat{\mu} &= \bar{y}_{\ldots} \\
\hat{\alpha}_i &= \bar{y}_{i..} - \bar{y}_{\ldots} \quad , i = 1,2,\ldots,a \\
\hat{\beta}_j &= \bar{y}_{.j.} - \bar{y}_{\ldots} \quad , j = 1,2,\ldots,b \\
(\hat{\alpha\beta})_{ij} &= \bar{y}_{ij.} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y}_{\ldots} \quad , \begin{cases} i = 1,2,\ldots,a \\ j = 1,2,\ldots,b \end{cases}
\end{aligned} \tag{43}$$

وإن الفرضيات التى يمكن اختبارها في تجارب من عاملين ثابتين تكون كالتالي :

1- $H_A : \alpha_1 = \alpha_2 = \ldots = \alpha_a = 0$

$H'_A :$ على الأقل إحدى قيم α_i لا تساوى صفر

2- $H_B : \beta_1 = \beta_2 = \ldots = \beta_b = 0$

$H'_B :$ على الأقل إحدى قيم β_j لا تساوى صفر

3- $H_{AB} : (\alpha\beta)_{ij} = 0\ \ , \forall (i,j)\ , i = 1,2,\ldots,a\ ,\ j = 1,2,\ldots,b$

$H'_{AB} :$ على الأقل إحدى قيم $(\alpha\beta)_{ij}$ لا تساوى صفر

كل اختبار من هذه الاختبارات سوف يكون مبنى على أساس مقارنة تقديرات مستقلة للتباين σ^2 ، هذه التقديرات يتم الحصول عليها من خلال تجزئة مجموع المربعات الكلى إلى أربعة عناصر مختلفة ، وذلك كما يلي :

$$\begin{aligned}
\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(y_{ijk} - \bar{y}_{ij.})^2 &= \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(y_{ijk} - \bar{y}_{ij.} + \bar{y}_{ij.} - \bar{y}_{i..} + \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y}_{.j.} - \bar{y}_{\ldots})^2 \\
&= \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{i=1}^{n}[(\bar{y}_{i..} - \bar{y}_{\ldots}) + (\bar{y}_{.j.} - \bar{y}_{\ldots}) + (\bar{y}_{ij.} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y}_{\ldots}) \\
&\quad + (y_{ijk} - \bar{y}_{ij.})]^2
\end{aligned}$$

$$= \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(\bar{y}_{i..} - \bar{y}_{...})^2 + \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(\bar{y}_{.j.} - \bar{y}_{...})^2$$
$$+ \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(\bar{y}_{ij.} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y}_{...})^2$$
$$+ \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(y_{ijk} - \bar{y}_{ij.})^2$$

جميع حدود التقاطع الضربى تساوى صفر وهى ست حدود ، فمثلاً

$$\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(\bar{y}_{i..} - \bar{y}_{...})(\bar{y}_{.j.} - \bar{y}_{...}) = n\sum_{i=1}^{a}(y_{i..} - \bar{y}_{...})\sum_{j=1}^{b}(\bar{y}_{.j.} - \bar{y}_{...}) = 0$$

ايضاً

$$\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(\bar{y}_{.j.} - \bar{y}_{...})(y_{ijk} - \bar{y}_{ij.}) = \sum_{i=1}^{a}\sum_{j=1}^{k}(\bar{y}_{.j.} - \bar{y}_{...})\sum_{k=1}^{n}(y_{ijk} - \bar{y}_{ij.})$$
$$= \sum_{i=1}^{a}\sum_{j=1}^{k}(\bar{y}_{.j.} - \bar{y}_{...})(n\bar{y}_{ij.} - n\bar{y}_{ij}) = 0$$

وهكذا بقية الحدود ، وعليه فإن :

$$\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(y_{ijk} - \bar{y}_{...})^2 = bn\sum_{i=1}^{a}(\bar{y}_{i..} - \bar{y}_{...})^2$$
$$+ an\sum_{j=1}^{b}(\bar{y}_{.j.} - \bar{y}_{...})^2$$
$$+ n\sum_{i=1}^{a}\sum_{j=1}^{b}(\bar{y}_{ij.} - \bar{y}_{i..} - \bar{y}_{.j.} + \dot{y}_{...})^2$$
$$+ \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(y_{ijk} - \bar{y}_{ij.})^2$$

اى أن :

$$SST = SSA + SSB + SSAB + SSE$$

حيث :

مجموع المربعات الكلى :

$$SST = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(y_{ijk} - \bar{y}_{...})^2 = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}y_{ijk}^2 - \frac{y^2}{abn} \qquad (44)$$

مجموع مربعات العامل A :

$$SSA = bn\sum_{i=1}^{a}(y_{i..} - \bar{y}_{...})^2 = \frac{\sum_{i=1}^{a} y_{i..}^2}{bn} - \frac{y_{...}^2}{abn} \quad (45)$$

مجموع مربعات العامل B :

$$SSB = an\sum_{j=1}^{b}(\bar{y}_{.j.} - \bar{y}_{...})^2 = \frac{\sum_{j=1}^{b} y_{.j.}^2}{an} - \frac{y_{...}^2}{abn} \quad (46)$$

مجموع مربعات التفاعل بين A و B :

$$SSAB = n\sum_{i=1}^{a}\sum_{j=1}^{b}(y_{ij.} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y}_{...})^2$$

$$= \frac{\sum_{i=1}^{a}\sum_{j=1}^{b} y_{ij.}^2}{n} - \frac{\sum_{i=1}^{a} y_{i..}^2}{bn} - \frac{\sum_{j=1}^{b} y_{.j.}^2}{an} + \frac{y_{...}^2}{abn} \quad (47)$$

مجموع مربعات الخطأ :

$$SSE = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(y_{ijk} - \bar{y}_{ij.})^2 = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n} y_{ijk}^2 - \frac{\sum_{i=1}^{a}\sum_{j=1}^{b} y_{ij.}^2}{n} \quad (48)$$

ومن الواضح أن SSA يقيس الاختلاف الكلى في متوسطات عينة العامل A عن $\bar{y}_{...}$ ، وإذا كان تأثير هذا العامل قليل أو ليس له تأثير على الإطلاق على μ فإن قيمة SSA ستكون صغيرة جداً ، أما SSB فإنه يقيس الاختلاف الكلى في متوسطات عينة العامل B عن $\bar{y}_{...}$ وبالمثل إذا كان ليس لهذا العامل تأثير أو أن تأثيره بسيط على μ فإن قيمته ستكون صغيرة أيضاً ، وإذا كان لا يوجد تفاعل بين العاملين فإن قيمة SSAB ستكون صغيرة ، ودرجات الحرية مجزئة بناء على المتطابقة $abn - 1 = (a - 1) + (b - 1) + (a - 1)(b - 1) + ab(n - 1)$ ، وبقسمة كل مجموع مربعات الذي بالطرف الأيمن من متطابقة مجموع المربعات على عدد درجات الحرية المناظرة له نحصل على أربعة مقدرات للتباين σ^2 وهي :

$$MSA = \frac{SSA}{a-1}, \quad MSB = \frac{SSB}{b-1}, \quad MSAB = \frac{SSAB}{(a-1)(b-1)}, \quad MSE = \frac{SSE}{ab(n-1)}$$

وكما في حالة تحليل التبـاين الأحـادي فإنـه يمكـن استخدام كل مجموع من مجاميع المربعات الأربعـة كتقدير للتبـاين σ^2 ، فـإذا نظرنـا لمجمـوع المربعـات كدالـة فـي المتغيرات العشـوائية المستقلة $Y_{111}, Y_{112}, \ldots, Y_{abn}$ فإنه يمكن الإثبات بأن :

$$E(MSA) = E\left[\frac{SSA}{a-1}\right] = \sigma^2 + nb\frac{\sum_{i=1}^{a}\alpha_i^2}{a-1}$$

$$E(MSB) = E\left[\frac{SSB}{b-1}\right] = \sigma_2 + \frac{na\sum_{j=1}^{b}\beta_j^2}{b-1}$$

$$E(MSAB) = E\left[\frac{SSAB}{(a-1)(b-1)}\right] = \sigma^2 + \frac{n\sum_{i=1}^{a}\sum_{j=1}^{b}(\alpha B)_{ij}^2}{(a-1)(b-1)}$$

$$E(MSE) = E\left[\frac{SSE}{ab(n-1)}\right] = \sigma^2$$

وعندمـا يكـون H_A صحيحـاً فـإن $E(MSA) = \sigma^2$ وعندمـا يكـون H_B صحيحـاً فـإن $E(MSB) = \sigma^2$ ، وعندما يكون H_{AB} فإن $E(MSAB) = \sigma^2$ ، ولكن إذا كان H_A خطأ فإن $E(MSA) > \sigma^2$ ، وبالتالي فإن : $F_A = \frac{MSA}{MSE}$ سيكون تقديراً للتناسب :

$$\frac{E(MSA)}{E(MSE)} = 1 + \frac{bn\sum_{i=1}^{a}\alpha_i^2}{(a-1)\sigma^2}$$

ويمكن الإثبات بأن : $F_A \sim f_{(\alpha,\ a-1,\ ab(n-1))}$ ،وعليه نرفض H_A إذا كانت :

$$F_A > f_{(\alpha,\ a-1,\ ab(n-1))} \quad .$$

وبالمثل في حالة H_B و H_{AB} سوف نرفض H_B إذا كانت :

$$F_B = \frac{MSB}{MSE} > f_{(\alpha,\ (b-1),\ ab(n-1))} \quad .$$

وسوف نرفض H_{AB} إذا كانت :

$$F_{AB} = \frac{MSAB}{MSE} > f_{(\alpha,\ (a-1)(b-1),\ ab(n-1))}$$

هذه النتائج يمكن تلخيصها في جدول تحليل تباين يطلق عليه تسمية جدول تحليل التباين الثنائي (2 - way Analysis of Variance) وذلك كما يلي :

جدول تحليل التباين الثنائي (2 - way ANOVA)

م.الاختلاف S.v	د.الحرية df	م.المربعات	متوسط المربعات MS	F المحسوبة
العامل A العامل B	a - 1 b - 1	SSA SSB	$MSA = SSA/(a-1)$ $MSB = SSB/(b-1)$	$F_A = MSA/MSE$ $F_B = MSB/MSE$
التفاعل AB الخطأ Error	$(a-1)(b-1)$ $ab(n-1)$	SSAB SSE	$MSAB = SSAB/(a-1)(b-1)$ $MSE = SSE/ab(n-1)$	$F_{AB} = MSAB/MSE$
الكلي Total	$abn - 1$	SST		

مثال (11) :لاختبار مدى فعالية ثلاث طرائق مختلفة للتدريس تم توزيع خمسة عشرة طالباً بطريقة عشوائية على ثلاث أساتذة وبطريقة عشوائية تم توزيع الطلاب على طرائق التعليم المختلفة ،أي خمسة طلاب لكل طريقة وتم تدريسهم نفس المنهج وبنهاية المدة المقررة لهذه التجربة أعطى نفس الاختبار لجميع الطلاب فكانت درجاتهم كما يلي :

طريقة التعليم (A)	الأستاذ (B)		
	I	II	III
1	75	90	85
	74	64	78
	86	80	77
	94	72	69
	78	68	82
2	92	90	82
	90	89	70
	84	91	73
	82	76	74
	75	84	83
3	85	69	89
	72	75	91
	80	98	73
	85	72	82
	76	73	84

والمطلوب تكوين جدول تحليل التباين واختبار في ما إذا كان للطرائق الثلاث نفس التأثير وإن للأساتذة الثلاث نفس الكفاءة ، ثم أختبر في ما إذا كان هناك تفاعل بين الطريقة والأستاذ عندما $\alpha = 0.05$ ؟

الحل :

لاختبار الفرضيات المطلوبة يجب إيجاد مجموع المربعات لجدول تحليل التباين وذلك كما يلي :

$$y_{...} = \sum_{i=1}^{3}\sum_{j=1}^{3}\sum_{k=1}^{5} y_{ijk} = 75 + 74 + \cdots + 69 + 82$$

$$+ 92 + 90 + \cdots + 83$$

$$+ 85 + 72 + \cdots + 82 + 84 = 3611$$

$$\sum_{i=1}^{3}\sum_{j=1}^{3}\sum_{k=1}^{5} y_{ijk}^2 = (75)^2 + (74)^2 + \cdots + (69)^2 + (82)^2$$

$$+ (92)^2 + (90)^2 + \cdots + (83)^2$$

$$+ (85)^2 + (72)^2 + \cdots + (82)^2 + (84)^2 = 292589$$

$$\bar{y}_{...} = \frac{y_{...}}{abn} = \frac{3611}{3 \times 3 \times 5} = 80.244$$

$$y_{11.} = \sum_{k=1}^{5} y_{11k} = 75 + 74 + 86 + 94 + 78 = 407 \Rightarrow \bar{y}_{11.} = \frac{y_{11}}{5} = \frac{407}{5} = 81.4$$

وهكذا بالنسبة لبقية القيم وعليه نتحصل على الجدول التالي :

طريقة التعليم (A)	الأستاذ (B) I	II	III	المجموع	المتوسط
1	$y_{11.} = 407$ $\bar{y}_{11.} = 81.4$	$y_{12.} = 374$ $\bar{y}_{12.} = 74.8$	$y_{13.} = 391$ $\bar{y}_{13.} = 78.2$	$y_{1..} = 1172$	$\bar{y}_{1..} = 78.133$
2	$y_{21.} = 423$ $\bar{y}_{21.} = 84.6$	$y_{22.} = 430$ $\bar{y}_{22.} = 86.0$	$y_{23.} = 382$ $\bar{y}_{23.} = 76.4$	$y_{2..} = 1235$	$\bar{y}_{2..} = 82.333$
3	$y_{31.} = 398$ $\bar{y}_{31.} = 79.6$	$y_{32.} = 387$ $\bar{y}_{32.} = 77.4$	$y_{33.} = 419$ $\bar{y}_{33.} = 83.8$	$y_{3..} = 1204$	$\bar{y}_{3..} = 80.267$
المجموع المتوسط	$y_{.1.} = 1228$ $\bar{y}_{.1.} = 81.867$	$y_{.2.} = 1191$ $\bar{y}_{.2.} = 79.4$	$y_{.3.} = 1192$ $\bar{y}_{.3.} = 79.467$	$y_{...} = 3611$	$\bar{y}_{...} = 80.244$

ومن هذا الجدول نجد أن :

$$SSA = \frac{\sum_{i=1}^{a} y_{i..}^2}{bn} - \frac{y_{...}^2}{abn} = \tfrac{1}{15}\{(1172)^2 + (1235)^2 + (1204)^2\} - \frac{(3611)^2}{45}$$
$$= 289895 - 289762.69 = 132.31$$

$$SSB = \frac{\sum_{j=1}^{b} y_{.j.}^2}{an} - \frac{y_{...}^2}{abn} = \tfrac{1}{15}\{(1228)^2 + (1191)^2 + (1192)^2\} - \frac{(3611)^2}{45}$$
$$= 289821.93 - 289762.69 = 59.24333$$

$$\frac{\sum_{i=1}^{n} \sum_{j=1}^{b} y_{ij.}^2}{n} = \tfrac{1}{5}\{(407)^2 \div (374)^2 + (391)^2 + (423)^2$$
$$+ (430)^2 + (384)^2 + (398)^2 + (387)^2 + (419)^2\} = 290378.6$$

$$SSAB = \frac{\sum_{i=1}^{a} \sum_{j=1}^{b} y_{ij.}^2}{n} - \frac{\sum_{i=1}^{a} y_{i..}^2}{bn} - \frac{\sum_{j=1}^{b} y_{.j.}^2}{an} + \frac{y_{...}^2}{abn}$$
$$= 290378.6 - 289895 - 289821.93 + 289762.69 = 424.36$$

$$SSE = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n} y_{ijk}^2 - \frac{\sum_{i=1}^{a}\sum_{j=1}^{b} y_{ij.}^2}{n} = 292589 - 290378.6 = 2210.40$$

$$SST = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n} y_{ijk}^2 - \frac{y^2}{abn} = 292589 - 289762.69 = 2826.31$$

ومن هذه الحسابات يكون جدول تحليل التباين (ANOVA) كما يلي :

جدول تحليل التباين

م.الاختلاف S.v.	د.الحرية df	م.المربعات	متوسط المربعات MS	F المحسوبة
العامل A	2	132.31	$MSA = 132.31/2 = 66.16$	$F_A = MSA/MSE = 1.078$
العامل B	2	59.24	$MSB = 59.24/2 = 29.62$	$F_B = MSB/MSE = 0.482$
التفاعل AB	4	424.36	$MSAB = 424.36/4 = 106.09$	$F_{AB} = MSAB/MSE = 1.728$
الخطأ Error	36	2210.40	$MSE = 2210.40/36 = 61.40$	
الكلي Total	44	2826.31		

ومن هذه النتائج يمكن إجراء اختبارات الفروض الآتية :

أ - اختبار في ما إذا كان لطرائق التعليم الثلاثة نفس التأثير :

$$H_A : \alpha_1 = \alpha_2 = \alpha_3 = 0$$

H'_A : على الأقل إحدى قيم α_i لا تساوى صفر

من جدول F وبدرجات حرية تساوى 2 و 36 و $\alpha = 0.05$ نجد أن :

$f_{(\alpha, a-1, ab(n-1))} = f_{0.05, 2, 36} = 3.275$ ، وحيث أن $F_A = 1.078 < 3.275$ وعليه لا توجد معلومات كافية لرفض H_A .

ب - اختبار في ما إذا كان للأساتذة الثلاثة نفس التأثير :

$$H_B : \beta_1 = \beta_2 = \beta_3 = 0$$

H'_B : على الأقل إحدى قيم β_j لا تساوى صفر

من جدول F وبدرجات حرية تساوى 2 و 36 و $\alpha = 0.05$ نجد أن :

$f_{(\alpha, b-1, ab(n-1))} = f_{0.05, 2, 36} = 3.275$ وحيث أن $F_B = 0.482 < 3.275$ ، وعليه لا توجد معلومات كافية لرفض H_B .

- اختبار في ما إذا كان هناك تفاعل ما بين الأستاذ والطريقة :

$H_{AB}: (\alpha\beta)_{ij} = 0 \ , \forall (i,j), i = 1,2,3 \ , \ j = 1,2,3$

على الأقل إحدى قيم $(\alpha\beta)_{ij}$ لا تساوي صفر : H'_{AB}

من جدول F وبدرجات حرية تساوي 4 و 36 و $\alpha = 0.05$ نجد أن :

$f_{(\alpha, (a-1)(b-1), ab(n-1))} = f_{0.05,4,36} = 2.65$ وحيث أن $F_{AB} = 1.729 < 2.65$ ، وعليه لا توجد معلومات كافية لرفض H_{AB} .

مثال (12) : تعتزم أمانة العدل شراء جهاز رادار يستخدم لتسجيل حالات الطوارئ التي تحدث على الطرقات ، وسوف تشترى الجهاز الذي شاشة عرضه أكثر كفاءة من حيث السرعة في تسجيل الحالة الطارئة ، ولهذا السبب قامت بتجريب ثلاث أجهزة في خمسة حالات طوارئ مختلفة ولإنجاز هذه التجربة تم توزيع 30 شخص متخصص في مجال المراقبة الجوية بطريقة عشوائية على الأجهزة والحالات المختلفة وسجل كل منهم وبشكل مستقل عن الأخريين الزمن (بالثواني) الذي يستغرقه كل جهاز لتسجيل الحالة الطارئة فكانت النتائج كما يلي :

الحالة الطارئة A	شاشة العرض B		
	1	2	3
1	18 16	13 15	24 28
2	31 35	33 30	42 46
3	22 27	44 41	40 37
4	39 36	35 38	52 57
5	15 12	10 16	28 24

والمطلوب :

أ - اختبار التفاعل بين العاملين .

ب - اختبار تساوي تأثير الحالات الطارئة .

د - اختبار تساوي تأثير شاشة العرض .

الحـل : هذه التجربة لها عاملين هما A وله 5 مستويات وهو يمثـل الحالـة الطارئـة و B ولـه 3 مسـتويات وهو يمثل شاشة العرض وبكل معالجة مفردتين (n=2) ، وعليه فإن عدد القيم يساوى 30 (N=abn =30) ، وسيكون مجموع الخلايا والمتوسطات ومجموع مربعات القيم كما يلي :

الحالة الطارئة	شاشة العرض 1	2	3	
1 المجموع	$y_{11.} = 34$	$y_{12.} = 28$	$y_{13.} = 52$	$y_{1..} = 114$
المتوسط	$\bar{y}_{11.} = 17$	$\bar{y}_{12.} = 14$	$\bar{y}_{13.} = 26$	$\bar{y}_{1..} = 19$
2 المجموع	$y_{21.} = 66$	$y_{22.} = 63$	$y_{23.} = 88$	$y_{2..} = 217$
المتوسط	$\bar{y}_{21.} = 33$	$\bar{y}_{22.} = 31.5$	$\bar{y}_{23.} = 44$	$\bar{y}_{2..} = 36.2$
3 المجموع	$y_{31.} = 49$	$y_{32.} = 85$	$y_{33.} = 77$	$y_{3..} = 211$
المتوسط	$\bar{y}_{31.} = 24.5$	$\bar{y}_{32.} = 42.5$	$\bar{y}_{33.} = 38.5$	$\bar{y}_{3..} = 35.2$

يتبع الجدول السابق :

4 المجموع المتوسط	$y_{41.} = 75$ $\bar{y}_{41.} = 37.5$	$y_{42.} = 73$ $\bar{y}_{42.} = 36.5$	$y_{43.} = 109$ $\bar{y}_{43.} = 54.5$	$y_{4..} = 257$ $\bar{y}_{4..} = 42.5$
5 المجموع المتوسط	$y_{51.} = 27$ $\bar{y}_{51.} = 13.5$	$y_{52.} = 26$ $\bar{y}_{52.} = 13$	$y_{53.} = 52$ $\bar{y}_{53.} = 26$	$y_{5..} = 105$ $\bar{y}_{5..} = 17.5$
المجموع المتوسط	$y_{.1.} = 251$ $\bar{y}_{.1.} = 25.1$	$y_{.2.} = 275$ $\bar{y}_{.2.} = 27.5$	$y_{.3.} = 378$ $\bar{y}_{.3.} = 37.8$	$y_{...} = 904$ $\bar{y}_{...} = 30.1$

ولن

$$\frac{\sum_{j=1}^{3} y_{.j.}^2}{10} = 28151 \quad , \frac{\sum_{i=1}^{5}\sum_{j=1}^{3} y_{ij.}^2}{2} = 31606 \quad , \sum_{i=1}^{5}\sum_{j=1}^{3}\sum_{k=1}^{2} y_{ijk}^2 = 31712 \quad , \frac{\sum_{i=1}^{5} y_{i..}^2}{6} = 30$$

ر $\frac{y_{...}^2}{30} = 27240533$ ، وعليه فإن :

$$SST = \sum_{i=1}^{5}\sum_{j=1}^{3}\sum_{k=1}^{2} y_{ijk}^2 - \frac{y_{...}^2}{abn} = 31712 - 27240.533 = 4471.467$$

$$SSA = \frac{\sum_{i=1}^{5} y_{i..}^2}{bn} - \frac{y_{...}^2}{N} = 30280 - 27240.533 = 3039.467$$

$$SSB = \frac{\sum_{j=1}^{3} y_{.j.}^2}{an} - \frac{y_{...}^2}{N} = 28151 - 27240.533 = 910.467$$

$$SSAB = \frac{\sum_{i=1}^{5}\sum_{j=1}^{3} y_{ij.}^2}{n} - \frac{\sum_{i=1}^{5} y_{i..}^2}{bn} - \frac{\sum_{j=1}^{3} y_{.j.}^2}{an} - \frac{y_{...}^2}{N}$$

$$= 31606 - 30280 - 28151 + 27240.533 = 415.533$$

$$SSE = \sum_{i=1}^{5}\sum_{j=1}^{3}\sum_{k=1}^{2} y_{ijk}^2 - \frac{\sum_{i=1}^{5}\sum_{j=1}^{3} y_{ij.}^2}{n} = 31712 - 31606 = 106$$

ومن هذه الحسابات يكون جدول تحليل التباين (ANOVA) كما يلي :

جدول تحليل التباين

م.الاختلاف S.V.	د.الحرية df	م.المربعات	متوسط المربعات MS	F المحسوبة
العامل A العامل B	4 2	3039.467 910.467	MSA = 3039.467/4 = 759.867 MSB = 910.467/2 = 455.234	$F_A = MSA/MSE = 107.523$ $F_B = MSB/MSE = 64.417$
التفاعل AB الخطأ Error	8 15	415.533 106.0	MSAB = 415.533/8 = 51.942 MSE = 106.0/15 = 7.067	$F_{AB} = MSAB/MSE = 7.350$
الكلي Total	29	4471.467		

ومن هذه النتائج يمكن اختبارات الفرضيات التالية :

أ - التفاعل :

$$H_{AB}:(\alpha\beta)_{ij} = 0 \ , \forall(i,j), i = 1,2,\ldots,5 \ , \ j = 1,2$$

H'_{AB} : على الأقل إحدى قيم $(\alpha\beta)_{ij}$ لا تساوى صفر

من جدول F وبدرجات حرية تساوى 8 و 15 وبمستوى معنوية يساوى 0.05 نجد أن $F_{AB} = 7.350 > f_{0.05,8,15} = 2.64$ ، وعليه نرفض H_{AB} ، أى أنه هناك مؤشر قوى على وجود التفاعل بين شاشة العرض والحالة الطارئة .

ب - تأثير الحالة الطارئة :

$$H_A : \alpha_1 = \alpha_2 = \ldots = \alpha_5 = 0$$

H'_A: على الأقل إحدى قيم α_i مختلفة

من جدول F وبدرجات حرية تساوى 4 و 15 وبمستوى معنوية يساوى 0.05 نجد أن $F_A = 107.523 > f_{0.05,4,15} = 3.06$ ، وعليه نرفض H_A .

جـ - تأثير شاشة العرض :

$$H_B : \beta_1 = \beta_2 = \beta_3 = 0$$

H'_B : مختلفة β_j على الأقل إحدى قيم

من جـدول F وبدرجـات حريـة تسـاوى 2 و 15 وبمسـتوى معنويـة يسـاوى 0.05 نجـد أن $F_B = 64.417 > f_{0.05,2,15} = 3.68$ ، وعليه نرفض H_B .

ملحوظة :

إذا وجد تفاعل في أى تجربة فإنه من الصعب التحدث عن التأثيرات الرئيسية للعاملين A و B ،وذلك لأنه متى وجد التفاعل سيكون تـأثير العـامل A متغير وذلـك لأنـه يعتمـد علـى مستوى العامل B ،وعليه فإن مقارنة متوسط تأثير العامل B (μ_j) عند مستويات مختلفة ليس له معنـى بالمثل بالنسبة لمقارنة متوسط تأثير (μ_i) العامل A عند مستويات مختلفة .

10 - 6 - 2 نموذج التأثير العشوائي The Random Effects Model

لقد تناولنا فيما سبق نموذج تحليل التباين الثنائي عندما كان العاملين A و B كلاهمـا ثـابت ، اي أن الباحث قد أختار مستويات معينة للعامل A وللعامل B .ولكن في هذا البند سـوف نتنـاول تحليل التباين الثنائي عندما تكون مستويات العاملين A و B قد تم اختيارهـا بطريقـة عشـوائية مـن مجتمعات كبيرة من المستويات، هـذا النمـوذج يسـمى نمـوذج التـأثير العشـوائي ،وإن أي اسـتنتاج حول هذه المستويات المختارة عشوائياً سوف يعمم على جميع مستويات المجتمعـات التـي سـحبت منها ، وسوف نفترض أن تأثيرات العامل A والعامل B والتي سنرمز لها بالرمز α_i و β_j على التوالي متغيرات عشـوائية مسـتقلة وأيضـاً تـأثيرات التفـاعل $(\alpha\beta)_{ij}$متغيرات عشـوائية مسـتقلة، ويمكن تمثيل مفردات هذا النموذج كما يلي :

$$y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ijk} \quad , \begin{cases} i = 1,2,\dots,a \\ j = 1,2,\dots,b \\ k = 1,2,\dots,n \end{cases} \qquad (49)$$

حيث :

$$\epsilon_{ijk} \sim NID(0,\sigma^2) \;\; \text{و} \;\; (\alpha\beta)_{ij} \sim NID(0,\sigma^2_{\alpha\beta})$$

$$\beta_j \sim NID(0,\sigma^2_\beta) \;\; \text{و} \;\; \alpha_i \sim NID(0,\sigma^2_\alpha) \quad \text{و}$$

وإن تباين أي مفردة سيكون كالآتي : $V(Y_{ijk}) = \sigma_\alpha^2 + \sigma_\beta^2 + \sigma_{\alpha\beta}^2 + \sigma^2$

إن تحليل التباين في حالة نموذج التأثيرات العشوائية لا يختلف عن تحليل التباين في حالة (50) تحليل نموذج التأثيرات الثابتة ، وبالتالي فإن الصيغ الرياضية لمجاميع المربعات بالنسبة إلى كل من SSA و SSBو SSAB و SSE و SSTستبقى كما هي بدون تغيير .ولكن الفرضيات الإحصائية ألتي سيتم اختبارها مختلفة عن السابق وذلك لأنها تتعلق بالتباين وليس بالمتوسط وبالتالي ستكون على النحو الآتي :

1- $H_A : \sigma_\alpha^2 = 0$ مقابل $H'_A : \sigma_\alpha^2 \neq 0$

2- $H_B : \sigma_B^2 = 0$ مقابل $H'_B : \sigma_B^2 \neq 0$

3- $H_{AB} : \sigma_{\alpha\beta}^2 = 0$ مقابل $H'_{AB} : \sigma_{\alpha\beta}^2 \neq 0$

وكما سبق ستكون إحصاءة الاختبار لأي فرضية من الفرضيات أعلاه مبنية على مقارنة متوسطي مجموعي مربعات لهما الخصائص الآتية :

أ - عندما يكون فرض العدم صحيحاً سيكون لهما نفس التوقع الرياضي .

ب - عندما يكون فرض العدم خطأ سيكون التوقع الرياضي لمتوسط المربعات الذي ببسط إحصاءة الاختبار أكبر من الذي في المقام .

ويمكن إثبات أن مثل تلك الإحصاءة تتبع في تغيراتها توزيع F عندما يكون فرض العدم صحيحاً، ولمعرفة الإحصاءة التي يجب استخدامها لاختبار أى فرضية من الفرضيات الثلاث أعلاه يجب معرفة التوقع الرياضي لمتوسط مجموع مربعات العامل A و B والتفاعل بينهما والأخطاء ، وسيكون هذا التوقع لكل منها كما يلي :

العامل A : $E(MSA) = \sigma^2 + bn\sigma_\alpha^2 + n\sigma_{\alpha\beta}^2$

العامل B : $E(MSB) = \sigma^2 + an\sigma_\beta^2 + n\sigma_{\alpha\beta}^2$

التفاعل AB : $E(MSAB) = \sigma^2 + n\sigma_{\alpha\beta}^2$

الخطاء : $E(MSE) = \sigma^2$

وبالرجوع للفرضيات أعلاه، نلحظ أنه لاختبار الفرضية (1) وبالنظر إلى MSA و MSAB نجد أن لهما نفس التوقع الرياضي عندما يكون فرض العدم صحيحاً اى عندما تكون $\sigma_\alpha^2 = 0$ أى أن العامل A ليس له تأثير ، ولكن عندما $\sigma_\alpha^2 \neq 0$ نجد أن E(MSA) أكبر من E(MSAB) وعليه فإن إحصاءة الاختبار المناسبة لهذه الفرضية ستكون كالآتي :

$$F_A = \frac{MSA}{MSAB}$$

هذه الإحصاءة تتوزع وفق توزيع F بدرجات حرية تساوى a - 1 و (a-1)(b-1) ، وعليه نرفض $H_A : \sigma_\alpha^2 = 0$ إذا كانت $F_A > f_{(\alpha, (a-1), (a-1)(b-1))}$.

وبالمثل لاختبار الفرضية (2) أى $H_B : \sigma_\beta^2 = 0$ ، سوف نستخدم إحصاءة الاختبار الآتية :

$$F_B = \frac{MSB}{MSAB}$$

وذلك لأنه عندما يكون فرض العدم صحيحاً سيكون التوقع الرياضي لكل من البسط والمقام يساوى $\sigma^2 + n\sigma_{\alpha\beta}^2$ ، ولكن عندما يكون خطأ ستكون $E(MSB) > E(MSAB)$. هذه الإحصاءة تتوزع وفق توزيع F بدرجات حرية تساوى b - 1 و (b-1)(a-1) ، وعليه نرفض $H_B : \sigma_\beta^2 = 0$ إذا كانت $F_B > f_{(\alpha, (b-1), (a-1)(b-1))}$.

أخيراً لاختبار الفرضية (3) : $H_{AB} : \sigma_{\alpha\beta}^2 = 0$ مقابل $H'_{AB} : \sigma_{\alpha\beta}^2 \neq 0$ وبالنظر إلى التوقع الرياضي لمتوسط مربعات الخطأ نجد أن إحصاءة الاختبار المناسبة تكون كالآتي :

$$F_{AB} = \frac{MSAB}{MSE}$$

وذلك لأنه عندما يكون فرض العدم صحيحاً سيكون التوقع الرياضي لكل من البسط والمقام يساوى σ^2 ، ولكن عندما يكون خطأ ستكون $E(MSAB) > E(MSE)$. إن التناسب F_{AB} سوف يتوزع وفق توزيع F بدرجات حرية تساوى ab (n-1) و (a-1)(b-1) ، وعليه نرفض فرض العدم إذا كانت $F_{AB} > f_{(\alpha, (a-1)(b-1), ab(n-1))}$.

مما سبق يتضح أن جميع الاختبارات أعلاه اختبارات من طرف واحد وإن إحصاءة الاختبار التى أستخدمت هنا تختلف عن التى استخدمناها عندما كان العاملان A و B ثابتين ، وبصفة عامة دائماً نستخدم التوقع الرياضي لمتوسط مربعات الخطأ كمنهاج عمل لاختيار إحصاءة الاختبار المناسبة .

ألحظ أنه يمكن تقدير σ_α^2 وذلك كما يلي :

$$E(MSA) - E(MSAB) = nb\sigma_\alpha^2 + n\sigma_{\alpha\beta}^2 + \sigma^2 - n\sigma_{\alpha\beta}^2 + \sigma^2 = nb\sigma_\alpha^2$$

وعليه فإن :

$$\hat{\sigma}_\alpha^2 = \frac{MSA - MSAB}{bn}$$

وبالمثل يمكن تقدير كل من σ_β^2 و $\sigma_{\alpha\beta}^2$ و σ^2 على التوالي كما يلي :

$$\hat{\sigma}_\beta^2 = \frac{MSB - MSAB}{an}$$

$$\hat{\sigma}_{\alpha\beta}^2 = \frac{MSAB - MSE}{n}$$

$$\hat{\sigma}^2 = MSE$$

مثال (13) :إذا افترضنا في المثال السابق أن مستويات العاملين A و B تم اختيارها بطريقة عشوائية ، أي أن كلا العاملين عشوائي فإن جدول تحليل التباين سيكون كما سبق عدا قيمة F المحسوبة ، وذلك يكون جدول تحليل التباين على النحو التالي :

جدول تحليل التباين

م.الاختلاف S.v.	د.الحرية df	م.المربعات	متوسط المربعات MS	F المحسوبة
العامل A العامل B	4 2	3039.467 910.467	MSA = 3039.467/4 = 759.867 MSB = 910.467/2 = 455.234	F_A = MSA/MSAB = 14.629 F_B = MSB/MSAB = 8.764
التفاعل AB الخطأ Error	8 15	415.533 106.0	MSAB = 415.533/8 = 51.942 MSE = 106.0/15 = 7.067	F_{AB} = MSAB/MSE = 7.350
الكلي Total	29	4471.467		

والفرضيات التي يمكن اختبارها هي :

1 - $H_A : \sigma_\alpha^2 = 0$ مقابل $H'_A : \sigma_\alpha^2 \neq 0$

حيث أن $F_A = 14.628 > f_{.05,4,8} = 3.84$ ، وعليه نرفض فرض العدم .

2 - $H_B : \sigma_B^2 = 0$ مقابل $H'_B : \sigma_B^2 \neq 0$

حيث أن $F_B = 8.764 > f_{.05,2,8} = 4.46$ ، وعليه نرفض فرض العدم .

3 - $H_{AB} : \sigma_{\alpha\beta}^2 = 0$ مقابل $H'_{AB} : \sigma_{\alpha\beta}^2 \neq 0$

حيث أن $F_{AB} = 3.44 < f_{(.05,8,15)} = 7.350$ ، وعليه لا توجد معلومات كافية لرفض فرض العدم .

ويمكن تقدير التباين كما يلي :

$$\hat{\sigma}_{\alpha}^{2} = \frac{MSA - MSAB}{bn} = \frac{759.867 - 51.942}{6} = 117.988$$

$$\hat{\sigma}_{\beta}^{2} = \frac{MSB - MSAB}{an} = \frac{455.234 - 51.942}{10} = 40.392$$

$$\hat{\sigma}_{\alpha\beta}^{2} = \frac{MSAB - MSE}{n} = \frac{51.942 - 7.067}{2} = 22.438$$

$$\hat{\sigma}^{2} = MSE = 7.067$$

10 - 6 - 3 النموذج المختلط The Mixed Model

في هذا البند سوف نتناول الحالة التي يكون فيها العامل A ذو تأثيرات ثابتة بينما مستويات العامل B قد تم اختيارها بطريقة عشوائية من مجتمع كبير من المستويات ، هذا النموذج يسمى بالنموذج المختلط ،وإن أى استنتاج حول المستويات المختارة عشوائياً سوف يعمم على جميع مستويات المجتمع الذي سحبت منه ، وسوف نرمز لتأثيرات العامل A والعامل B بالرمز α_i و β_j على التوالي ، ولتأثيرات التفاعل بينهما بالرمز $(\alpha\beta)_{ij}$ ، ويمكن تمثيل مفردات هذا النموذج كما يلي :

$$y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \in_{ijk} \qquad , \begin{cases} i = 1,2,\dots,a \\ j = 1,2,\dots,b \\ k = 1,2,\dots,n \end{cases} \qquad (51)$$

حيث α_i تأثير ثابت بحيث $\sum_{i=1}^{a} \alpha_i = 0$ ، وإن

$$\in_{ijk} \sim NID(0,\sigma^2) \text{ و } (\alpha\beta)_{ij} \sim NID(0,\frac{a-1}{a}\sigma_{\alpha\beta}^{2}) \text{ و } \beta_j \sim NID(0,\sigma_{\beta}^{2})$$

ألحظ أن أي حدي تفاعل $(\alpha\beta)_{ij}$ و $(\alpha\beta)_{kl}$ تكونا مستقلين إلا إذا كان كلاهما يشير لنفس المستوى من العامل العشوائي B ، وفي مثل تلك الحالة فإن التغاير بينهما يساوى $-\frac{1}{a}\sigma_{\alpha\beta}^{2}$ ،

وإن $\sum_i (\alpha\beta)_{ij} = 0$ لجميع قيم j وذلك لأنه يتضمن جميع مستويات العامل A بينما $\sum_j (\alpha\beta)_{ij}$ لا يساوى صفر .

وإن تباين أي مفردة سيكون كالأتي :

$$V(Y_{ijk}) = \sigma_\beta^2 + \sigma_{\alpha\beta}^2 + \sigma^2$$

إن تحليل التباين في حالة النموذج المختلط لا يختلف عن تحليل التباين في حالة تحليل نموذج التأثيرات العشوائية أو الثابتة ، وبالتالي فإن الصيغ الرياضية لمجاميع المربعات بالنسبة إلى كل من SSA و SSB و SSAB و SSE و SST ستبقى كما هي بدون تغيير .إن الفرضيات الإحصائية التي سيتم اختبارها هنا تكون على النحو الآتي :

-1

$$H_A : \alpha_i = 0$$

H'_A : على الأقل واحدة من α_i لا تساوي صفر

-2

$$H_B : \sigma_B^2 = 0$$
$$H'_B : \sigma_B^2 \neq 0$$

-3

$$H_{AB} : \sigma_{\alpha\beta}^2 = 0$$
$$H'_{AB} : \sigma_{\alpha\beta}^2 \neq 0$$

وكما سبق ستكون إحصاءة الاختبار لأي فرضية من الفرضيات أعلاه مبنية على مقارنة متوسطي مجموعي مربعات لهما الخواص الآتية :

أ – عندما يكون فرض العدم صحيحاً سيكون لهما نفس التوقع الرياضي .

ب – عندما يكون فرض العدم خطأ سيكون التوقع الرياضي لمتوسط المربعات الذي ببسط إحصاءة الاختبار أكبر من الذي في المقام .

ويمكن إثبات أن مثل تلك الإحصاءة تتبع في تغيراتها توزيع F عندما يكون فرض العدم صحيحاً، ولمعرفة الإحصاءة التي يجب استخدامها لاختبار أي فرضية من الفرضيات الثلاث أعلاه يجب معرفة التوقع الرياضي لمتوسط مجموع مربعات العامل A و B والتفاعل بينهما والأخطاء ، وسيكون هذا التوقع لكل منها كما يلي :

العامل A : $$E(MSA) = \sigma^2 + n\sigma^2_{\alpha\beta} + \frac{bn\sum_{i=1}^{a}\alpha_i^2}{a-1}$$

العامل B : $$E(MSB) = \sigma^2 + an\sigma^2_{\beta}$$

التفاعل AB : $$E(MSAB) = \sigma^2 + n\sigma^2_{\alpha\beta}$$

الخطاء : $$E(MSE) = \sigma^2$$

بالرجوع للفرضيات أعلاه، نلحظ أنه لاختبار الفرضية (1) وبالنظر إلى MSA وMSAB نجد أن لهما نفس التوقع الرياضي عندما يكون فرض العدم صحيحاً أى عندما تكون $\alpha_i = 0$ ، ولكن عندما لا تساوى جميع α_i الصفر نجد أن $E(MSA)$ أكبر من $E(MSAB)$ وعليه فإن إحصاءة الاختبار المناسبة لهذه الفرضية ستكون كالآتي :

$$F_A = \frac{MSA}{MSAB}$$

هذه الإحصاءة تتوزع وفق توزيع F بدرجات حرية تساوى a - 1 و (b-1)(a-1) ، وعليه نرفض $H_A : \alpha_i = 0$ إذا كانت $F_A > f_{(\alpha\,,(a-1),(a-1)(b-1))}$.

وبالمثل لاختبار (2) أى $H_B : \sigma^2_\beta = 0$ ، سوف نستخدم إحصاءة الاختبار الآتية :

$$F_B = \frac{MSB}{MSE}$$

وذلك لأنه عندما يكون فرض العدم صحيحاً سيكون التوقع الرياضي لكل من البسط والمقام يساوى σ^2 ، ولكن عندما يكون خطأ ستكون $E(MSB) > E(MSE)$.هذه الإحصاءة تتوزع وفق توزيع F بدرجات حرية تساوى b - 1 و ab(n-1) ،وعليه نرفض $H_B : \sigma^2_\beta = 0$ إذا كانت $F_B > f_{(\alpha\,,(b-1),an(b-1))}$.

وأخيراً لاختبار الفرضية (3) : $H_{AB} : \sigma^2_{\alpha\beta} = 0$ مقابل $H'_{AB} : \sigma^2_{\alpha\beta} \neq 0$ وبالنظر إلى التوقع الرياضي لمتوسط مربعات الخطأ نجد أن إحصاءة الاختبار المناسبة تكون كالآتي :

$$F_{AB} = \frac{MSAB}{MSE}$$

وذلك لأنه عندما يكون فرض العدم صحيحاً سيكون التوقع الرياضي لكل من البسط والمقام يساوى σ^2 ، ولكن عندما يكون خطأ ستكون $E(MSAB) > E(MSE)$.إن التناسب F_{AB} سوف يتوزع وفق توزيع F بدرجات حرية تساوى (b-1)(a-1) و ab (n-1) ،وعليه نرفض فرض العدم إذا كانت $F_{AB} > f_{(\alpha\,,(a-1)(b-1),ab(n-1))}$.

وكما أشرنا سابقاً سوف نستخدم التوقع الرياضي لمتوسط مربعات الخطأ كمنهاج عمل لاختبار إحصاءة الاختبار المناسبة ويمكن تقدير كل من σ_β^2 و $\sigma_{\alpha\beta}^2$ و σ^2 على التوالي كمايلي:

$$\hat{\sigma}_\beta^2 = \frac{MSB - MSE}{an}$$

$$\hat{\sigma}_{\alpha\beta}^2 = \frac{MSAB - MSE}{n}$$

$$\hat{\sigma}^2 = MSE$$

مثال (14) : في بيانات المثال (12) وعلى افتراض أن الباحث قد اختار مستويات معينة للعامل A ، بينما مستويات العامل B قد تم اختيارها بطريقة عشوائية من مجتمعها ، أختبر صحة الفرضيات التالية :

أ - $H_A : \alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = \alpha_5 = 0$

$H'_A :$ على الأقل واحدة من α_i لا تساوي صفر

ب - $H_B : \sigma_B^2 = 0$ مقابل $H'_B : \sigma_B^2 \neq 0$

جـ - $H_{AB} : \sigma_{\alpha\beta}^2 = 0$ مقابل $H'_{AB} : \sigma_{\alpha\beta}^2 \neq 0$

الحـل :

لاختبار هذه الفرضيات نكون جدول تحليل التباين وسيكون كما سبق مع مراعاة التغير في قيمة F المحسوبة .

جدول تحليل التباين الثنائي للنموذج المختلط

(2-way ANOVA for mixed effects model)

م.الاختلاف S.v	د.الحرية df	م.المربعات	متوسط المربعات MS	F المحسوبة
العامل A	4	3039.467	$MSA = 3039.467/4 = 759.867$	$F_A = MSA/MSAB = 14.629$
العامل B	2	910.467	$MSB = 910.467/2 = 455.234$	$F_B = MSB/MSE = 64.417$
العامل AB	8	415.533	$MSAB = 415.533/8 = 51.942$	$F_{AB} = MSAB/MSE = 7.350$
الخطأ Error	15	106.0	$MSE = 106.0/15 = 7.067$	
الكلي Total	29	4171.467		

ا - باستخدام جدول F وبدرجات حرية تساوى 4 و 15 وبمستوى معنوية يساوى 0.05 نجد أن $f_{0.05,4,15} = 3.06 < F_A = 107.417$ ، وعليه نرفض H_A .

ب - من جدول F وبدرجات حرية تساوى 2 و 15 وبمستوى معنوية يساوى 0.05 نجد أن $f_{0.05,2,15} = 3.68 < F_B = 64.417$ ، وعليه نرفض H_B .

جـ - من جدول F وبدرجات حرية تساوى 8 و 15 وبمستوى معنوية يساوى 0.05 نجد أن $f_{0.05,8,15} = 2.64 < F_{AB} = 7.350$ ، وعليه نرفض H_{AB} .

تمرينات Exercises

1 - الجدول التالي يبين تحليل التباين لتجربة ثنائيه :

مصدر الأختلاف	درجات الحرية	مجموع المربعات	متوسط المربعات	F المحسوبة
A B AB Error	3 5	2.6 9.2 ... 18.7	... ... 3.1	... ...
المجموع	47			

أ - أكمل جدول تحليل التباين .

ب - كم عدد مستويات كل عامل وعدد المفردات بكل توفيق .

جـ - هل يوجد تفاعل بين العاملين عند مستوى المعنوية 5 ٪ ؟

2 - أجريت دراسة على خمسة أنواع من السيارات وذلك لغرض اختبار تساوي متوسط المسافة التي يقطعها كل نوع بكل جالون بنزين ، حيث تم اختيار عينة عشوائية مستقلة من كل نوع وكل سيارة تمت قيادتها حتى أستنفدت كمية البنزين بالكامل علماً بأنه بخزان كل سيارة جالون واحد فقط ، فكانت المسافة التي قطعتها كل سيارة كما يلي :

النوع	المسافة بالميل
A :	18 ، 17 ، 18 ، 21 ، 19
B :	21 ، 24 ، 17 ، 23 ، 22 ، 23
C :	15 ، 14 ، 16 ، 15
D :	18 ، 20 ، 20 ، 24 ، 26 ، 23 ، 25
E :	17 ، 16 ، 18 ، 17 ، 15 ، 17

من هذه البيانات هل يمكن القول بأن الأنواع الخمسة تختلف في كمية البنزين المستهلك بكل ميل.

3 ...

مستوى مفردتين :

العامل B	العامل A		
	1	2	3
1	2 4	5 6	1 3
2	5 4	2 2	10 9
3	7 10	1 0	5 3
4	8 7	12 11	7 4

والمطلوب :

أ – تكوين جدول تحليل التباين .

ب – أختبر الفرضيات المناسبة في الحالات الآتية :

1 – عندما يكون النموذج ثابت ، 2 – عندما يكون العامل A ثابت والعامل B متغير .

جـ – هل يوجد تفاعل بين العاملين عند مستوى المعنوية 1 ٪ ؟

4 – قام أحد الأساتذة بقسم الحاسب الآلي بتصميم استبيان لقياس كفاءة المتقدمين للعمل بقسم الحاسب ، وطلب من كل متقدم كتابة مستوى معلوماته وخبرته في الحاسب (A = ممتاز B = متوسط C = دون المتوسط) قبل البدء في الإجابة وتم اختيار عينة عشوائية من المستويات الثلاث فكانت درجاتهم كما يلي (علماً بأن درجة الامتحان من 150) :

A : 80 , 90 , 93 , 82 , 114 , 88 , 80 , 105

B : 130 , 133 , 110 , 130 , 90 , 104 ,128

C : 151 , 140 ,156 , 128

والمطلوب

أ – كون جدول تحليل التباين لهذه التجربة .

ب – هل هناك فروق بين متوسطات المستويات الثلاث عند مستوى المعنوية 5 ٪ ؟

جـ – أختبر وجود فروق معنوية بين جميع الأزواج الممكنة عند مستوى المعنوية 5 ٪ .

د – تحقق من أن MSE هو مقدر التباين المشترك (σ^2) وذلك من خلال الأثبات بأن :

$$\cdot\ MSE = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2 + (n_3 - 1)s_3^2}{(n_1 - 1) + (n_2 - 1) + (n_3 - 1)}$$

5 - البيانات التالية تمثل عدد علب السجائر المباعة من 4 أنواع من السجائر خلال ثمانية أيام تم اختيارهم بطريقة عشوائية :

A : 45 , 60 , 33 , 36 , 31 , 40 , 43 , 48
M : 35 , 12 , 27 , 41 , 19 , 23 , 31 , 20
R : 21, 35 , 32 , 28 , 14 , 47 , 25 , 38
S : 32 , 53 , 29 , 42 , 40 , 23 , 35 , 42

والمطلوب :

أ - كون جدول تحليل التباين .

ب - هل هناك فروق معنوية بين الأنواع الأربعة عند مستوى المعنوية 5 ٪ ؟

جـ - أستخدم اختبار توكي وفيشر لتحديد أي المتوسطات تختلف عن بعضها البعض .

د - قارن : I - متوسط R ومتوسط S مع متوسط M .

II - متوسط A ومتوسط M مع متوسط R ومتوسط S .

III - متوسط M ومتوسط R ومتوسط S مع متوسط A .

جـ - أختبر تجانس تباين المجتمعات الأربعة .

6 - صُممت تجربة لمقارنة درجة تآكل أربعة أنواع من الإطارات بعد استعمالها مسافة قدرها 34000 كم بطريقة التصميم العشوائي الكامل بقطاعات ، حيث استخدمت أربعة سيارات كقطاعات وكل أربعة اطارات تم تركيبها في كل سيارة بترتيب عشوائي ثم تمت قيادة كل سيارة 34000 كم ، فكانت نتائج التآكل في كل نوع من الإطارات كما يلي :

السيارة	نوع الإطار			
	A	B	C	D
1	15	11	10	11
2	12	10	10	9
3	11	10	8	10
4	11	8	8	8

من هذه البيانات هل يمكن القول بأنه توجد فروق معنوية في متوسط درجة التآكل بين الأنواع الأربعة من الإطارات عند مستوى المعنوية 1 % ؟

7 – البيانات التالية تمث عدد الوحدات المنتجة من قبل ثلاثة عمال حيث كل منهم يعمل على نفس الآلة ولمدة أربعة أيام مختلفة

الآلة	العامل		
	A	B	C
1	19 ، 18 ، 18 ، 19	18 ، 23 ، 22 ، 23	19 ، 16 ، 20 ، 21
2	18 ، 16 ، 17 ، 17	16 ، 16 ، 15 ، 15	18 ، 19 ، 16 ، 18
3	16 ، 17 ، 18 ، 15	18 ، 18 ، 17 ، 18	15 ، 16 ، 15 ، 17

اختبر عند مستوى المعنوية 5 % اذا كان

أ- هناك فروق بين الآلات الثلاث .

ب- هناك فروق بين العمال الثلاث .

جـ- هناك تفاعل بين العمال والآلات .

8 – في تجربة لعلم البيولوجيا استخدم 3 أنواع من التركيز الكيميائي التي تساعد في نمو نوع معين من النباتات خلال فترة زمنية معينة والبيانات التالية تمثل قياسات الطول المأخوذة من النباتات التي عاشت خلال فترة التجربة :

التركيز		
I	II	III
8.0	7.4	7.1
7.7	6.9	6.8
8.1	6.8	6.9
8.4	5.8	7.3
8.6	7.2	6.3
9.4	8.7	6.1

والمطلوب :

أ- تكوين جدول تحليل التباين ، ثم عند مستوى المعنوية 1 ٪ اختبار الفرضيات المناسبة في الحالات التالية :

I - نموذج التأثير الثابت . II - نموذج التأثير العشوائي .

ب - استخدم اختبار توكي لتحديد أي المتوسطات تختلف عن بعضها البعض عند مستوى المعنوية 1 ٪ .

9 - البيانات التالية تمثل المتوسط التراكمي لثلاثين طالب حسب الجنس والتخصص

الجنس		التخصص
ذكور	إناث	
3.8 ، 4.0 ، 3.9	3.0 ، 2.6 ، 2.8	رياضيات
2.6 ، 3.0 ، 2.9	3.9 ، 3.4 ، 3.6	احصاء
4.0 ، 3.6 ، 3.9	2.7 ، 3.2 ، 3.8	فيزياء
3.4 ، 3.6 ، 3.4	3.5 ، 3.7 ، 3.8	كيمياء
3.5 ، 3.8 ، 3.7	4.0 ، 3.5 ، 3.8	نبات

كون جدول تحليل التباين الثنائي ثم أختبر الفرضيات المناسبة في الحالات الأتية :

أ - العاملين ثابتين . ب - التخصص ثابت والجنس عشوائي .

جـ - التخصص عشوائي والجنس ثابت . د - كلاهما عشوائي .

10- صممت تجربة بطريقة التصميم العشوائي الكامل بقطاعات وذلك لغرض المقارنة بين ثلاثة معالجات A ، B ، C في أربعة قطاعات فكانت النتائج كما يلي :

المعالجة	القطاع			
	1	2	3	4
A	3	6	2	1
B	5	7	6	4
C	3	3	2	2

هل هناك فرق بين تأثير المعالجات الثلاثة عندما $\alpha = 0.5$ ، أوجد 90 ٪ فترة ثقة حول $(\mu_A - \mu_B)$.

11 – البيانات التالية تمثل الدرجات النهائية لأربعة طلاب في 4 مقررات :

الطالب	المقرر M	S	E	B
1	78	62	71	77
2	71	66	59	67
3	57	49	62	60
4	69	78	72	83

من هذه البيانات وعند مستوى المعنوية 10 ٪ ، أختبر الفرضيات التالية :

أ – المقررات الأربعة متساوية من حيث الصعوبة .

ب – للطلاب الأربعة نفس الكفاءة .

12– لكي نقارن بين أربعة أنواع من طرق التدريس لمادة الحاسوب من حيث التحصيل العلمي أختيرت عينة عشوائية مستقلة من مجموعات كبيرة من الطلاب الذين تم تدريسهم بالطرق الأربعة ، فكانت درجاتهم في امتحان عام كما يلي :

الطريقة	الدرجات
A :	75 ، 73 ، 68 ، 72
B :	84 ، 92 ، 84 ، 82 ، 87 ، 85 ، 87
C :	62 ، 65 ، 68 ، 67 ، 67 ، 66
D :	74 ، 76 ، 73 ، 72 ، 76 ، 74 ، 75 ، 79

من هذه البيانات هل يمكن القول بأن هناك فروق معنوية بين طرق التدريس الأربعة .

13– أربعة أنواع من اللحوم تم اختيارهم وذلك لغرض معرفة كمية الدهون الموجودة بها ، ومن كل نوع تم اختيار عينة عشوائية فكانت النتائج كما يلي :

النوع	كمية الدهون (٪)
1	41 ، 42 ، 40 ، 44 ، 43
2	38 ، 34 ، 36 ، 37 ، 38 ، 36
3	42 ، 45 ، 48 ، 46 ، 47 ، 48
4	54 ، 52 ، 51 ، 52 ، 53

هل هناك فروق معنوية من حيث متوسط كمية الدهون بالأنواع الأربعة عند مستوى المعنوية 5 ٪؟

14- يوجد ثلاثة مصارف في مدينة مصراتة ومن كل مصرف تم اختيار الزبائن بطرائق عشوائية وسجل لكل منهم الزمن الذي ينتظره حتى تقدم له خدمة فكانت النتائج كما يلي :

المصرف	الفترة الزمنية المنتظرة
A :	12.5 ، 13.0 ، 13.5 ، 16.0 ، 13.5 ، 14.5
B :	14.5 ، 18.0 ، 16.5 ، 16.0 ، 17.5 ، 16.0 ، 17.5
C :	12.5 ، 13.5 ، 14.0 ، 13.5 ، 12.5

من هذه البيانات هل هناك فروق معنوية بين المصارف الثلاثة من حيث متوسط الزمن الذي ينتظره الزبائن حتى تقدم لهم الخدمة . أي المتوسط تختلف عن بعضها البعض .

15 - تحليل التباين لتصميم عشوائي كامل بقطاعات أعطى النتائج التالية :

المصدر S.V	درجات الحرية	مجموع المربعات	متوسط المربعات	F المحسوبة
TRT.	3	27.1	...	...
Block	5	...	14.90	...
Error	...	33.4	...	
المجموع	...			

أ - أكمل الجدول .

ب - هل هذه البيانات تشير لوجود فروق معنوية بين تأثير المعالجات عندما $\alpha = 0.01$.

ج - اذا كانت $\bar{y}_A = 9.7$ و $\bar{y}_B = 12.1$ أوجد 90 ٪ فترة ثقة حول $(\mu_A - \mu_B)$.

16- دواء لعلاج الجلوكوما تم اختباره على عشرة كلاب مريضة ، الدواء تم تطبيقه على عين واحدة يتم اختيارها بطريقة عشوائية لكل كلب ، وبعد ساعة من العلاج يتم قياس الضغط على كرات العينين لكل كلب ، وعليه فإن التجربة تتضمن معالجتين ، الأولى داء الجلوكوما ، والثانية لاعلاج (مراقبة أو تحكم) والكلاب العشرة تمت معاملتها كقطاعات ، وكانت نتائج قياسات الضغط (القياس الأكبر يعني شدة الإصابة بالمرض) كما يلي :

الكلب		1	2	3	4	5	6	7	8	9	10
المعالجات	علاج	0.15	0.18	0.13	0.18	0.19	0.12	0.07	0.09	0.14	0.06
	لاعلاج	0.17	0.20	0.14	0.18	0.29	0.19	0.12	0.10	0.16	0.13

أ - كون جدول تحليل التباين ثم اختبر الفرضية أنه لا يوجد فرق بين المعالجتين .

ب - ما الهدف من عمل الكلاب كقطاعات بهذه التجربة .

جـ - حلل هذه البيانات باستخدام t للبيانات المزدوجة ثم قارن قيمة t المحسوبة مع قيمة F المحسوبة ثم تحقق من أن $F_{\alpha,1,n-1} = t^2_{\frac{\alpha}{2},n-1}$ ، وبالتالي عندما يكون عدد المعالجات يساوي 2 فإن اختبار F للتصميم العشوائي الكامل بقطاعات يكافئ اختبار t للبيانات المزدوجة .

17- إن أحد الأمور المهمة التي تحدد موقع انشاء مجمع تجاري جديد هو عدد السيارات المارة بذلك الموقع طيلة اليوم ولهذا قام أحد رجال الأعمال بوضع أربعة أشخاص كعددين في أربعة مواقع مختلفة وحصر عدد السيارات التي تمر خلال كل موقع ولمدة خمسة أيام فكانت النتائج كما يلي :

اليوم	الموقع			
	1	2	3	4
1	226	241	222	197
2	250	302	252	245
3	196	200	181	195
4	220	225	214	202
5	214	216	220	215

أ - ما هو التصميم المناسب لهذه التجربة .

ب - هل هناك فروق معنوية بين متوسط عدد السيارات المارة كل يوم بالمواقع الأربعة .

الفصـــل الحادي عشــر
الإحصاء اللاّمعلمى
Nonparametric Statistics

11-1 مقدمـــة Intoduction

إن معظم الأساليب التــي أتبعناها فـي اختبـارات الفروض وتكوين فترات الثقة في الفصـول السابقة مبنية على الفرضية التي تقول بأن العينة أو العينات العشوائية التي تم اختيارها للدراسة هي من مجتمع (أو مجتمعات) طبيعي، وغالباً ما تكون هذه الأســاليب غير حساسة إلى حد ما عندما يكون مجتمع العينة (أو العينات) غير طبيعي، وحيث أن بعض المجتمعات لا تفي بالشرط المطلوب لتطبيق تلك الأساليب ، دعت الحاجة للبحث عن أساليب أخـرى لا يتطلب تطبيقها مثل تلك الشرط ، هذه الأساليب يطلق عليها تسمية الأساليب اللامعلمية لأنه وكما لاحظنا في الفصـول السابقة سواء عند التقدير أو اختبارات الفروض كان اهتمامنـا منصب علـى معلمة أو أكثر من معلمات المجتمع الاحصائي (المتوسط ، التبـاين ، النسبة ، ... الخ) عـلاوة علـى ذلك ، وكما أشرنا في تلك الفصول لكي نصل إلى استنتاج إحصائي يجب معرفة صيغة التوزيـع الاحتمـالي للمجتمع (أو المجتمعات) التي تم اختيار العينة (أو العينات) منها.

إن الأساليب التي تهتم بدراسة معلمــة (أو معلمـات) المجتمـع الإحصـائي يطلـق عليها تسمية أساليب معلميه ، وهناك نوعان من الأساليب الإحصـائية تتم معاملتهما على أنها أسـاليب لامعلميـه وهما : أساليب لامعلميه بما تعنيه الكلمة وهي أساليب تختبر الفرضيات التي لا تتضمـن أي نـص يتعلق بمعلمات المجتمع الإحصائي ، أما الأساليب الأخرى فهي أسـاليب التوزيعـات الحـرة وهـى الأساليب التي لا تضع أي افتراضات على مجتمـع المعاينـة ، وبصـرف النظـر عن التمييز بين هذان الأسلوبان فإن كلاهما ستتم معاملتها على أنهما أساليب لامعلميه ، هذه الأساليب يتم تطبيقهـا على سبيل المثال لا الحصر في الحالات الآتية :

1- إذا كانت الفرضية المطلوب اختبارها لا تتضمن معلمة المجتمع .

2- البيانات مقاسه بمقياس أضعف من المقـاييس المطلوبـة لتطبيق الأسـاليب المعلميـة ، وسـوف نبين فيما بعد ما هو المقصـود بهذه المقاييس .

3- إن لم تتوفر الشروط المطلوبة لتطبيق الأساليب المعلميه .

ومن مزايا الأساليب اللامعلميه ما يلي :

. ضعيف بمقياس مقاسه البيانات تكون عندما تطبيقها يمكن - أ

ويمكن إنجازها بسرعة . الحسابات الضرورية للأساليب اللامعلميه عادةً ما تكون سهلة - ب

أ - يمكن تطبيقها عندما تكون البيانات مقاسه بمقياس ضعيف .

ب - تعتمد على افتراضات قليلة ، وبالتالي فرصة تطبيقها خطأ ستكون صغيرة .

جـ - الحسابات الضرورية للأساليب اللامعلميه عادةً ما تكون سهلة ويمكن إنجازها بسرعة .

د - سهولة فهمها وطريقة حسابها تجعلها مناسبة جداً للباحثين الذي ليس لديهم خلفية علمية جيدة في الرياضة والإحصاء .

بالرغم من المزايا السابقة للأساليب اللامعلميه إلا أنه يعيب عليها نتيجة لسهولة حسابها ، في بعض الأحيان يتم تطبيقها في مسائل يكون من الأفضل تطبيق أساليب معلميه عليها مما يسبب في ضياع المعلومات .

لقد استعملنا في الفقرات السابقة كلمة المقياس والتي يمكن تعريفها على أنها طريقة تخصيص الأعداد للمشاهدات أو الأحداث قيد الدراسة بناءً على مجموعة من القواعد ، وبالتالي اختلاف مجموعة القواعد المتبعة في إعطاء الأعداد للمشاهدات يولد عنها مقاييس مختلفة هذه المقاييس هي:

أ - المقياس الأسمي : The nominal scale

يميز هذا المقياس القياسات أو الصفات أو القراءات المأخوذة عن ظاهرة ما عن بعضها البعض من خلال إعطائها تسمية معينة ، فمثلاً يمكن تصنيف الإنتاج من سلعة ما على أنه قابل للاستهلاك أو غير قابل للاستهلاك ، أو مثلاً تصنيف المولود على أنه ذكر أو أنثى .... وهكذا ، وعادةً ما تستخدم أرقام اختيارية للتمييز بدلاً من الاسم ، فمثلاً يمكن إعطاء العدد " 1 " إذا كان الإنتاج غير قابل للاستهلاك والعدد " 0 " إذا كان الإنتاج قابل للاستهلاك، وعادةً ما يستخدم هذا المقياس إذا كان الهدف من الدراسة معرفة المشاهدات التي تقع في التصنيفات الاسمية المختلفة وهو من أضعف المقاييس الإحصائية .

ب - المقياس الترتيبي : Ordinal scale

عندما تكون المشاهدات ليست مختلفة من صنف إلى صنف فقط ولكن يمكن ترتيبها أيضاً بناءً على معيار معين عندئذ يقال أنها مقاسه بمقياس ترتيبي ، فمثلاً تصنيف ذكاء طالب على أنه دون المتوسط أو متوسط أو أعلى من متوسط ففي هذه الحالة العناصر الموجودة في كل تصنيف متساوية ولكن عناصر أي تصنيف يمكن اعتبارها على أنها أفضل أو أسوأ من التصنيف الآخر ، وعليه يتضح أن المقياس الترتيبي يجعل من الممكن إعطاء رتب للمشاهدات وليس بالضرورة أن تكون الفروق ما بين الرتب متساوية .

ج- المقياس الفتروى : Interval scale

عندما يمكن تمييز المفردات عن بعضها البعض ويمكن ترتيب الفروق ما بين أي قياسين له معنى فإنه يمكن تطبيق القياس الفتروى فمثلاً إذا أعطيت الدرجات 5 ، 10 ، 20 ، 25 للطلاب A ، B ، C ، D فإنه يمكن القول بأن الفرق ما بين 5 و 10 يساوى الفرق ما بين 20 و 25 إذن بخلاف القياس الأسمى والترتيبى فإن القياس الفتروى هو قياس كمي . ومن الأمثلة المألوفة على المقياس الفتروى هو مقياس درجة الحرارة بالدرجات الفهرنهايتية والمئوية حيث قراءة الدرجة صفر في الترمومتر بالفهرنهايت أو بالمئوي لا يعني عدم وجود درجة حرارة .

د- القياس النسبي : The ratio scale

عندما تكون للقياسات خواص المقاييس الثلاثة السابقة بالإضافة إلى أن التناسب ما بين هذه القياسات له معنى عندئذٍ يقال بأن القياس قياس نسبى ومن الأمثلة على القياس النسبي قياسات الأوزان والأطوال ويعتبر هذا المقياس من أقوى المقاييس .

من الناحية العملية إذا كان حجم العينة أقل من 50 وتباين المجتمع غير معروف وقمنا برسم المضلع التكراري وكان غير معتدلاً ، أو إذا أمكن حساب معامل الالتواء وكانت قيمة هذا المعامل كبيرة فإنه يفضل استخدام الأساليب اللامعلميه .

إن هذه الأساليب أصبحت في العقود الأخيرة تمثل مجالاً من مجالات علم الإحصاء وبالتالي فإن الاختبارات التي سنتعرض إليها في هذا الفصل ما هي إلا جزء بسيط من هذا المجال .

2-11 اختبار الإشارة The sign test

يعتبر هذا الاختبار من أقدم الاختبارات اللامعلميه حيث يرجع آلي سنة (1710م) ونظراً لسهولة استخدامه فإن أوجه استعمالاته عديدة ، ويعتبر هذا الاختبار مفيد على وجه الخصوص عندما تكون وحدة القياس ترتيبي ، حيث يمكن المقارنة خلال كل زوج من المفردات وتحديد أي المفردات أكبر من الأخرى داخل الزوج الواحد . وهناك العديد من الحالات التي يمكن فيها تطبيق هذا الاختبار ، يمكن أيضاً تطبيق اختبارات لا معلميه أكثر قوة ، وكما سنرى في ما بعد إن هذا الاختبار يحول البيانات قيد الدراسة إلى سلسلة من الإشارات الموجبة والسالبة ، ومن هنا كانت تسميته .

شروط تطبيق الاختبار : عينة عشوائية ثنائية (X_i , Y_i) , $i = 1,2,3,...,n$ حيث كل

1- تتضمن البيانات قياسات من عينة عشوائية ثنائية من نفس المفردة أو المفردات التي ازدوجت بالنسبة زوج من هذه القياسات تم الحصول عليه من نفس المفردة أو المفردات التي ازدوجت بالنسبة لمتغير او أكثر (فمثلاً من الممكن أن تكون X_i عدد دقات قلب مريض قبل تعاطى الدواء بينما Y_i تمثل عدد دقات قلبه بعد تعاطى الدواء)،وخلال كل زوج (X_i , Y_i) تتم المقارنة كما يلي :

$$(X_i , Y_i) = \begin{cases} + & , \text{إذا كانت} \quad X_i > Y_i \\ - & , \text{إذا كانت} \quad X_i < Y_i \\ 0 & , \text{إذا كانت} \quad X_i = Y_i \end{cases}$$

ويجب أن تحذف الأزواج التي تساوى صفر من التحليل ويتناقص حجم العينة (n) تبعاً لذلك .

2 - يجب أن تكون القياسات (X_i , Y_i) , $i = 1,2,3,...,n$ مستقلة عن بعضها البعض .

3 - يجب أن يكون المتغير قيد الدراسة متغيراً متصلاً .

4 - وحدة القياس على الأقل ترتيبي ، بحيث يمكن تحديد أي المقياسين أكبر خلال كل زوج .

الفرضيات :

أ - اختبار من طرفين :

$$H_0 : P(+) = P(-)$$
$$H_1 : P(+) \neq P(-)$$

ب - اختبار من طرف واحد :

$$H_0 : P(+) \leq P(-)$$
$$H_1 : P(+) > P(-)$$

جـ - اختبار من طرف واحد :

$$H_0 : P(+) \geq P(-)$$
$$H_1 : P(+) < P(-)$$

ألحظ أنه يمكن صياغة الفرضيات أعلاه بدلالة الوسيط فمثلاً في الحالة (أ) يمكن صياغة H_0 و H_1 على النحو الآتي :

H_0 : وسيط مجتمع الفروق $(X_i - Y_i)$ يساوى صفر

H_1 : وسيط مجتمع الفروق $(X_i - Y_i)$ لا يساوى صفر

وبالمثل يمكن صياغتها في حالة اختبار من طرف واحد .

مثلاً في حالة (ب) :

وسيط مجتمع الفروق $(X_i - Y_i)$ أصغر من أو يساوى صفر : H_0

وسيط مجتمع الفروق $(X_i - Y_i)$ أكبر من صفر : H_1

إحصاءة الاختبار :

1 - في الحالة (أ) : إذا كان H_0 صحيحاً فإننا نتوقع بأن يكون عدد الإشارات الموجبة " + " مساوياً لعدد الإشارات السالبة " - " ، وعليه إذا كان عدد أي منهما صغيراً فسيؤدى ذلك إلى رفض H_0 ، وبالتالي فإن إحصاءة الاختبار (T) ستكون مساوية لعدد الإشارات الموجبة أو عدد الإشارات السالبة أيهما أقل .

2 - في الحالة (ب) : إن ما يتضمنه H_0 في هذه الحالة هو أن قيم X_i تميل إلى أن تكون أصغر من قيم Y_i ، وعليه إذا كان عدد الإشارات السالبة صغيراً سيؤدى ذلك إلى رفض H_0 وبالتالي فإن : T = عدد الإشارات السالبة .

3 - في الحالة (جـ) : في هذه الحالة فإن H_0 يشير إلى أن قيم X_i تميل إلى أن تكون أكبر من قيم Y_i ، وعليه إذا كان عدد الإشارات الموجبة صغيراً سيؤدى ذلك إلى رفض H_0 وبالتالي فإن : T = عدد الإشارات الموجبة .

القرار :

من جدول (2) وبمعلومية $\tilde{p} = \frac{1}{2}$ و n = عدد الإشارات الموجبة + عدد الإشارات السالبة ، ومستوى المعنوية (α) سيكون القرار كما يلي :

1 - في حالة اختبار من طرفين : نرفض H_0 عند مستوى المعنوية α إذا كان

$$P(T \le t | n , \tilde{p} = \frac{1}{2}) \le \frac{\alpha}{2}$$

حيث t تمثل القيمة المشاهدة لإحصاءة الاختبار T و T متغير عشوائي (عدد الإشارات في حالة H_0) يتبع توزيع ذي الحدين بمعلمتين هما n و $\tilde{p} = \frac{1}{2}$.

2 - في حالة اختبار من طرف واحد : سواء كانت الحالة (ب) أو الحالة (جـ) فإننا نرفض H_0 عند مستوى المعنوية α إذا كان

$$P(T \le t \mid n\ ,\ \tilde{p} = \frac{1}{2}) < \alpha$$

حيث t تمثل القيمة المشاهدة لإحصاءة الاختبار T في الحالة (2) أو (3) .

ملحوظة :

I - إن اختبار الإشارة لاختبار عينتين ذات علاقة يمكن أيضاً استخدامه كاختبار للوسيط إذا كانت البيانات تمثل عينة عشوائية $X_1, X_2, X_3, \cdots, X_n$ من مجتمع وسيطه (m) غير معروف. وإن الفرضيات الآتية مناظرة للفرضيات السابقة .

أ - اختبار من طرفين :

$$H_0 : m = m_0$$
$$H_1 : m \neq m_0$$

حيث m_0 تمثل قيمة الوسيط الفرضية.

ب - اختبار من طرف واحد :

$$H_0 : m \le m_0$$
$$H_1 : m > m_0$$

جـ - اختبار من طرف واحد :

$$H_0 : m \ge m_0$$
$$H_1 : m < m_0$$

ولحساب إحصاءة الاختبار تطرح قيمة الوسيط الفرضية (m_0) من كل قيمة من قيم العينة وتسجل إشارة الفرق ، أي تسجل إشارات الفروق $X_i - m_0\ ,\ i = 1,2,3,\cdots,n$ ثم نتبع نفس الخطوات من (1) إلى (3) السابقة وسيكون القرار كما هو بدون تغيير .

II - إذا كان حجم العينة كبيراً وكانت np و $n(1 - p)$ أكبر من 5 ، فإنه يمكن استخدام التقريب الطبيعي لحساب قيمة t وذلك كما يلي :

في حالة اختبار من طرفين :

$$t = \frac{1}{2}(n + z_{\frac{\alpha}{2}} \sqrt{n})$$

وفي حالة اختبار من طرف واحد فإن :

$$t = \frac{1}{2}(n + z_{\alpha}\sqrt{n})$$

مثال (I) : قام ستة طلبة بإتباع نظام معين في الأكل وذلك كمحاولة لتخفيف أوزانهم فكانت النتائج كما يلي :

الطالـــــــــب	1	2	3	4	5	6
الوزن قبل تنظيم الأكل (X)	80	75	77	84	90	70
الوزن بعد تنظيم الأكل (Y)	75	65	81	72	80	66

وعلى ضوء هذه البيانات هل يمكن القول بأن تنظيم الأكل كان له دور في تخفيف الوزن عند مستوى المعنوية 5٪ .

الحل :

الفرضية :

إن ما يتضمنه الفرض البديل H_1 في هذه الحالة هو أن الأوزان قبل تنظيم الأكل تميل إلى أن تكون أكبر من الأوزان بعد التنظيم ، وعليه فإن

$$H_0 : P(+) \leq P(-)$$

$$H_1 : P(+) > P(-)$$

إحصاءة الاختبار : لحساب إحصاءة الاختبار تطرح الأوزان بعد تنظيم الأكل من الأوزان قبل تنظيمه وذلك كما يلي :

الطالـــــــــب	1	2	3	4	5	6
الوزن قبل تنظيم الأكل (X)	80	75	77	84	90	70
الوزن بعد تنظيم الأكل (Y)	75	65	81	72	80	66
$X_i - Y_i$	5	10	4-	12	10	4
إشارة الفرق	+	+	−	+	+	+

وعليه فإن T= عدد الإشارات السالبة =1

القرار :

حيث أن $n = 5 + 1 = 6$ و $\tilde{p} = \frac{1}{2}$ وإن $\alpha = 0.05$ ، وبالتالي من جدول (2) نجد أن

$$P(T \leq 1 | n = 6, \tilde{p} = \frac{1}{2}) = 0.1094$$

حيث أن 0.1094 أكبر من α ، وعليه لا توجد معلومات كافية لرفض H_0 عند مستوى المعنوية 0.05 ، ومستوى المعنوية المشاهد (p - value) يساوى 0.1094 .

مثال (2) : الجدول الآتى يعرض نتائج الفروق فى فعالية 12 زوج من المستحضرات الصيدلية ثم تحليلها باستخدام طريقتين هما X و Y .

المستحضر	1	2	3	4	5	6	7	8	9	10	11	12
الفروق $X_i - Y_i$	-.5	-.4	.7	-.3	-.1	-.6	.2	.3	-.4	-.8	-.2	.9

والهدف هو اختبار عدم وجود فرق ما بين الطريقتين عند مستوى المعنوية 5 % .

الحـل :

الفرضية :

$$H_0 : P(+) = P(-)$$
$$H_1 : P(+) \neq P(-)$$

إحصاءة الاختبار : T = عدد الإشارات الموجبة أو السالبة أيهما أقل ، وعليه فإن

المستحضر	1	2	3	4	5	6	7	8	9	10	11	12
إشارة الفرق :	-	-	+	-	-	-	+	+	-	-	-	+

حيث أن عدد الإشارات الموجبة أقل من عدد الإشارات السالبة وعليه فإن T = 4.

القرار :

حيث أن $n = 4 + 8 = 12$ و $\tilde{p} = \frac{1}{2}$ وإن $\alpha = 0.05$ نجد أن

$$P(T \le 4 \mid n = 12, \tilde{p} = \frac{1}{2}) = 0.1937$$

وحيث أن 0.1937 أكبر من $\frac{\alpha}{2} = 0.025$ ، وعليه لا توجد معلومات كافية لرفض H_0 ، ومستوى المعنوية المشاهد يساوى 0.1937 .

مثال (3) : بفرض أن البيانات الآتية تمثل عينة من أوزان مواليد بأحد المستشفيات

المولود	1	2	3	4	5	6	7	8	9	10	11
الوزن X_i	3.3	2.25	1.8	3.25	2.75	3.1	3.5	2.7	3	4	2.5

هل يمكن القول بأن وسيط أوزان المواليد بالمجتمع الذي اختيرت منه العينة يختلف عن 3.5 كجم عند مستوى المعنوية 5 ٪ .

الحل :

الفرضية :

وسيط أوزان المواليد بالمجتمع يساوى 3.5 كجم $H_0 : (m = 3.5)$

وسيط أوزان المواليد بالمجتمع لا يساوى 3.5 كجم $H_1 : (m \ne 3.5)$

إحصاءة الاختبار :

لحساب إحصاءة الاختبار نوجد الفروق $X_i - 3.5$ وذلك كما يلي :

المولود	1	2	3	4	5	6	7	8	9	10	11
الوزن X_i	3.3	2.25	1.8	3.25	2.75	3.1	3.5	2.7	3	4	2.5
$X_i - 3.5$	−	−	−	−	−	−	0	−	−	+	−

حيث أنه هناك تطابق ما بين الوسيط وإحدى القيم وعليه تهمل هذه المفردة من الدراسة وبالتالي فإن :

n = عدد الإشارات الموجبة + عدد الإشارات السالبة = 1 + 9 = 10

وحيث أن عدد الإشارات السالبة أكبر من عدد الإشارات الموجبة، وعليه فإن إحصاءة الاختبار تساوى عدد الإشارات الموجبة أى أن $T = 1$.

القرار :

ومن جدول (2) ، وبمعلومية $n = 10$ و $\tilde{p} = \frac{1}{2}$ نجد أن

$$P(T \leq 1 | n = 10 , \tilde{p} = \frac{1}{2}) = 0.0108$$

وحيث أن 0.0108 أصغر من $\frac{\alpha}{2} = 0.025$ (لأن الاختبار من طرفين) ، وعليه نرفض H_0 عند مستوى المعنوية 5 ٪ ، أى أن وسيط أوزان المواليد بالمجتمع الذى أختيرت منه العينة يختلف عن 3.5 كجم ، ومستوى المعنوية المشاهد (p-value) يساوى $2 (0.0108) = 0.0216$ وهو أقل من 0.05 .

3-11 اختبار رتب الإشارة ولكاكسن The wilcoxon signed ranks test

لقد تعرضنا في ما سبق لاختبار الإشارة ، وأشرنا بالقول إلى أن المعلومة الوحيدة التى يستخدمها هذا الاختبار عند تحليل البيانات المزدوجة (X_i , Y_i) هو تحديد ما إذا كانت المفردة X_i أكبر أو أصغر أو أنها تساوى المفردة Y_i ، وهو من أفضل الاختبارات التي تستخدم في مثل هذه النوعية من البيانات وخاصة إذا كانت وحدة قياسها ضعيفة ، أما إذا كانت وحدة قياسها قوية فإن استخدامه قد يؤدى لفقدان بعض المعلومات التي تتضمنها البيانات ، وبالتالي ضعف قوة الاستنتاج الإحصائي ، وعليه سنتعرض الآن لاختبار أخر يستخدم معظم المعلومات التي تتضمنها البيانات المزدوجة ألا وهو اختبار ولكاكسن .

إن هذا الاختبار يفضل استخدامه عندما يمكننا تحديد مقدار الفرق الموجود بين أى زوجين من المفردات (X_i , Y_i) بالإضافة إلى اتجاه ذلك الفرق ، وعندما يكون بالإمكان تحديد مقدار تلك الفروق ، فإنه يمكن ترتيب هذه الفروق ومن هذا الترتيب فإن الاختبار يستخدم أكبر قدر ممكن من المعلومات التي تضمنتها البيانات الأصلية ، أى أن هذا الاختبار يحول المفردتان اللتان بالزوج (X_i , Y_i) إلى مفردة واحدة وذلك من خلال دراسة الفرق :

$$D_i = Y_i - X_i \quad , i = 1 , 2 , \cdots , n$$

وبالتالي فإن التحليل سيجرى باستخدام قيم D_i كعينة من المفردات ، وملاحظة حجم قيم D_i الموجبة مقارنة بقيم D_i السالبة ، إن أهم فرق ما بين اختبار الإشارة واختبار ولكاكسن هو أن

هذا الأخير يشترط أن يكون توزيع مجتمع هذه الفروق متماثل ، وعلى ضوء هذا الشرط فإن أى استنتاج إحصائي يتعلق بالوسيط سوف يكون صحيح بالنسبة للمتوسط ، وإن بُعد أى مفردة عن الوسيط سيكون له معنى ، وبالتالي فإن مقدار البُعد بين أى مفردتين سيكون مقياس ذو معنى . وبناء على ذلك فإن وحدة القياس المطلوبة سيكون قياس فترة (Interval) .

إن اختبار ولكاكسن مُعد لمعرفة ما إذا كانت بيانات العينة قيد الدراسة ثم اختيارها من مجتمع له وسيط معين ، بالإضافة إلى إمكانية استخدامه في الحالات التي يمكن الحصول منها على قراءات أو قياسات قبل تعريض مفرداتها لاختبار معين ثم بعد تعريضها للاختبار ، وذلك لمعرفة هل أن المتغير الثاني في الزوج له نفس الوسيط مثل المتغير الأول .

شروط تطبيق الاختبار :

1- تحتوى البيانات على n من القيم للفروق $D_i = Y_i - X_i$ ، $i = 1, 2, \cdots, n$ وإن كل زوج من القياسات (X_i , Y_i) ثم الحصول عليه من نفس الوحدة التجريبية أو وحدات تجريبية ازدوجت بالنسبة لمتغير أو أكثر .

2- توزيع مجتمع الفروق D_i , $i = 1,2,...,n$ توزيع متماثل .

3- يجب أن تكون الفروق D_i , $i = 1,2,...,n$ مستقلة عن بعضها البعض .

4- جميع الفروق D_i , $i = 1,2,...,n$ لها نفس الوسيط .

5- وحدة قياس الفروق D_i على الأقل فتروى .

الفرضيات :

إذا كانت $\tilde{\mu}_D$ ترمز لوسيط مجتمع الفروق (D_i) فإنه يمكن كتابة الفرضيات كما يلي :

أ - اختبار من طرفين :

$H_0 \; : \; (\tilde{\mu}_D = 0)$ وسيط مجتمع الفروق يساوى صفر

$H_1 \; : (\tilde{\mu}_D \neq 0)$ وسيط مجتمع الفروق لا يساوى صفر

ب - اختبار من طرف واحد :

$H_0 \; : (\tilde{\mu}_D \leq 0)$ وسيط مجتمع الفروق أصغر من أو يساوى صفر

$H_1 \; : (\tilde{\mu}_D > 0)$ وسيط مجتمع الفروق أكبر من الصفر

جـ - اختبار من طرف واحد :

$H_0 : (\tilde{\mu}_D \geq 0)$: وسيط مجتمع الفروق أكبر من أو يساوى صفر

$H_1 : (\tilde{\mu}_D < 0)$: وسيط مجتمع الفروق أصغر من الصفر

إحصاءة الاختبار :

لحساب قيمة إحصاءة الاختبار نتبع الخطوات الآتية :

1- إيجاد الفروق $D_i = Y_i - X_i$, $i = 1, 2, \ldots, n$ وإذا كانت أى من هذه الفروق تساوى صفر تهمل من الدراسة ويتناقص حجم العينة تبعاً لذلك .

2- ترتيب الفروق المطلقة أى $|D_i|$, $i = 1, 2, \ldots, n$ من الأصغر إلى الأكبر ، وإذا تساوت قيمتين أو أكثر من قيم $|D_i|$ سيعطى لها متوسط رتبها ، فمثلاً إذا كانت أصغر أربعة فروق مطلقة متساوية فإننا نرتبها كالآتي : 1 ، 2 ، 3 ، 4 ثم تعطى لكل منها الرتبة 2.5 (أى أن $\frac{4+3+2+1}{4} = \frac{10}{4}$) .

3- تعطى الرتب الناتجة إشارة الفرق المناظر لها .

4- إيجاد مجموع الرتب التي إشارتها موجبة (T_+) وعليه فإن إحصاءة الاختبار تكون كالآتي :

أ - إذا كان لا يوجد مفردات متساوية أو أن عددها قليل جداً فإن احصاءة الاختبار تكون (T_+) .

ب -إذا كان هناك عدد كبير من المفردات متساوية فإن احصاءة الاختبار تكون كالآتي :

$$T = \frac{\sum_{i=1}^{n} R_i}{\sqrt{\sum_{i=1}^{n} R_i^2}}$$

حيث R_i ترمز لرتب D_i الموجبة والسالبة .

القرار :

إذا كانت ω_α ترمز للتجزئ ذو المرتبة α والمتحصل عليه من جدول (9) عند استخدام T_+ أو من جدول (4) عند استخدام T (مع مراعاة استبدال T بدلاً من T_+ في الآتي) وإن القرار سوف يعتمد على الفرضيه قيد الاختبار وذلك على النحو التالي :

1- في الحالة (أ) : الإحصاءة تكون T_+ ، وبمعلومية n و α نوجد $\omega_{\frac{\alpha}{2}}$ ونرفض H_0 إذا كانت $T<\omega_{\frac{\alpha}{2}}$ أو $T>\omega_{1-\frac{\alpha}{2}}$ ، حيث $\omega_{\frac{\alpha}{2}}$ تمثل القيمة الجدولية بجدول (9) .

من الناحية التطبيقية وعندما تكون قيم X_i تمثل القياسات عن ظاهرة معينة قبل معالجتها و Y_i تمثل القياسات بعد المعالجة فإن رفض H_0 يعني أن المعالجة لها تأثير أما إذا كانت البيانات تمثل نتائج تم الحصول عليها من خلال مقارنة معالجتين مختلفتين فإن رفض H_0 يعني أن المعالجتين لهما تأثير مختلف .

2- في الحالة (ب) :- إذا كانت قيمة T_+ كبيرة فإنها مؤشر على عدم صحة فرض العدم وعليه نرفض H_0 إذا كانت قيمة $T_+ > \omega_{1-\alpha}$ ، وبالتالي إذا كانت البيانات تتعلق بنتائج مقارنة معالجتين مختلفتين ، فإن رفض H_0 يعني أن أحد المعالجتين لها تأثير أكثر من الأخرى .

3- في الحالة (جـ) :- إذا كانت قيمة T_+ صغيرة فإنها تدل على عدم صحة H_0 وعليه نرفض H_0 إذا كانت قيمة $T_+ < \omega_\alpha$ ، وبالتالي إذا كانت البيانات تتعلق بنتائج مقارنة معالجتين مختلفتين ، فإن رفض H_0 يعني أن أحد المعالجتين لها تأثير أكثر من الأخرى .

ملحوظة :

إن اختبار رتب الإشارة لولكاكسن من الممكن استخدامه أيضاً لاختبار الوسيط عندما تكون البيانات قيد الدراسة تم الحصول عليها من عينة عشوائية واحدة ، فإذا كانت $X_1, X_2, X_3, \ldots, X_n$ تمثل عينة من مجتمع وسيطه ($\tilde{\mu}$) غير معروف ، وكانت $\tilde{\mu}_0$ تمثل قيمة الوسيط الفرضية وهو مقدار ثابت فإن الفرضيات الآتية مناظرة للفرضيات السابقة .

أ - اختبار من طرفين :

$$H_0 : \tilde{\mu} = \tilde{\mu}_0$$
$$H_1 : \tilde{\mu} \neq \tilde{\mu}_0$$

ب - اختبار من طرف واحد :

$$H_0 : \tilde{\mu} \geq \tilde{\mu}_0$$
$$H_1 : \tilde{\mu} < \tilde{\mu}_0$$

جـ - اختبار من طرف واحد :

$$H_0 : \tilde{\mu} \leq \tilde{\mu}_0$$
$$H_1 : \tilde{\mu} > \tilde{\mu}_0$$

ولإيجاد إحصاءة الاختبار يطرح من كل مفردة من المفردات (X_i) المقدار $\tilde{\mu}_0$ ، أى ان الفروق تكون كالآتي :

$$D_i = \tilde{\mu}_0 - X_i \quad , i = 1,2,3,\ldots,n$$

وإتباع الخطوات من (1) إلى(4) التي سبق وأن أشرنا إليها في كيفية إيجاد إحصاءة الاختبار ، وستبقى القاعدة أيضاً كما هي بدون أى تغيير ، وحيث أن التوزيع متماثل ، وبالتالي يمكن استبدال كلمة الوسيط بالمتوسط .

مثال (4) : بفرض أن البيانات الآتية تمثل عينة من أعمار الطلبة الذين التحقوا بالسنة الأولى بالجامعة خلال 16 سنة ماضية :

18.5 ، 19 ، 20 ، 17.4 ، 17.9 ، 18 ، 17.6 ، 17

18.1 ، 19.4 ، 20.5 ، 21 ، 19.3 ، 17.2 ، 18.8 ، 19.3

هل يمكن القول بأن وسيط أعمار مجتمع هذه العينة يختلف عن 18 عند مستوى المعنوية 0.05 .

الحل :

الفرضية :

H_0 : وسيط مجتمع العينة يساوى 18

H_1 : وسيط مجتمع العينة لا يساوى 18

إحصاءة الاختبار : لحساب الاحصاءة تكون الجدول الآتي :

العمر	$D_i = 18 - X_i$	رتبة $\lvert D_i \rvert$	R_i
17	1	8.5	+8.5
17.6	0.4	3	+3
18	0	–	–
17.9	0.1	1.5	+1.5
17.4	0.6	5	+5
20	−2	13	-13
19	−1	8.5	-8.5
18.5	−0.5	4	-4
19.3	−1.3	10.5	-10.5
18.8	−0.8	6.5	−6.5
17.2	0.8	6.5	6.5
19.3	−1.3	10.5	−10.5
21	−3	15	−15
20.5	−2.5	14	−14
19.4	−1.4	12	−12
18.1	−0.1	1.5	−1.5

حيث إن مجموع الرتب الموجبة $(T_+) = 24.5$ ، وبالتالي إحصاءة الاختبار هي $T_+ = 24.5$.

القرار :

سوف نرفض H_0 إذا كانت $T < \omega_{\frac{\alpha}{2}}$ أو $T > \omega_{1-\frac{\alpha}{2}}$ ، وحيث أن $n = 15$ و $\frac{\alpha}{2} = 0.025$ وعليه فإن $\omega_{0.025} = 26$ ، وبالتالي نرفض H_0 ومستوى المعنوية المشاهد $0.02 < p - value < 0.05$.

مثال (5) : بفرض أن البيانات الآتية تمثل معدلات نبضات القلب لعينة من المرضى قبل وبعد إجراء عملية جراحية :

المريض	قبل العملية X_i	بعد العملية Y_i
1	69	72
2	75	73
3	68	78
4	71	81
5	73	70
6	77	75
7	70	83
8	65	74
9	60	75
10	74	70

هل هذه البيانات تشير إلى أن معدل نبضات القلب يزداد بعد العملية الجراحية عند مستوى المعنوية 5٪ ؟

الحل :

الفرضية :

وسيط مجتمع الفروق أصغر من أو يساوى صفر $H_0 : (\tilde{\mu}_D \leq 0)$

وسيط مجتمع الفروق أكبر من الصفر $H_1 : (\tilde{\mu}_D > 0)$

إحصاءة الاختبار :

المريض	x_i	y_i	$D_i = y_i - x_i$	رتبة $\|D_i\|$	R_i
1	69	72	3	3.5	3.5
2	75	73	-2	1.5	-1.5
3	68	78	10	7.5	7.5
4	71	81	10	7.5	7.5
5	73	70	-3	3.5	-3.5
6	77	75	-2	1.5	-1.5
7	70	83	13	9	9
8	65	74	9	6	6
9	60	75	15	10	10
10	74	70	-4	5	-5

ستكون مجموع الرتب الموجبة أى أن $T_+ = 43.5$

القرار :

من جدول (9) و $n = 10$ و $\alpha = 0.05$ نجد أن $\omega_{0.05} = 11$ وإن

$$\omega_{0.95} = \frac{n(n+1)}{2} - \omega_{0.05} = 55 - 11 = 44$$

وحيث أن $44 > T_{+}$ ، وعليه لا توجد معلومات كافية لرفض H_0 ومستوى المعنوية المشاهد أكبر من 0.05 وأقل من 0.10 .

11- 4 اختبار مان- وايتني The mann - whitney test

يستخدم هذا الاختبار لاختبار الفرضية H_0 التي تهدف إلى معرفة مدى تطابق مجتمعين من حيث معلمتي الموقع (المتوسط أو الوسيط) ، وذلك على أساس اختيار عينتين عشوائيتين منهما على أن تكون بيانات العينتين من نوع ترتيبي (ordinal) أي أنه يساعد في الإجابة على الأسئلة التي من النوع :"هل أحد المجتمعين يبدو أنه يعطي قيم أكبر من المجتمع الآخر ؟" أو " هل وسيطي المجتمعين متساويين ؟" . ويعتبر هذا الاختبار من أقوى الاختبارات اللامعلمية المستخدمة لهذا الغرض ، ويستخدم هذا الاختبار رتب المفردات بدلاً من المفردات نفسها ويفضل استخدام الرتب للأسباب الآتية :

أولاً : إذا كانت الأعداد المعطاة للمفردات لا معنى لها بحد ذاتها ولكن يكون لها معنى في حالة مقارنتها بالترتيب مع الأعداد الأخرى فقط أي أن الأعداد لا تحتوي على معلومات أكثر مما تحتويه الرتب وهذا من طبيعة البيانات التي من نوع ترتيبي .

ثانياً : حتى إذا كان لهذه الأعداد معنى ولكن دالة التوزيع لا تتبع التوزيع الطبيعي ، فإن نظرية الاحتمالات عادةً لا تكون في متناولنا عندما تكون إحصاءة الاختبار تعتمد على البيانات الحقيقية ، علاوة على ذلك إن نظرية الاحتمالات المبنية على الرتب تعتبر نسبياً سهلة ولا تعتمد على التوزيع في كثير من الحالات .

ثالثاً : إن الكفاءة النسبية لاختبار مان - وايتني ليست بسيئة مقارنة باختبار t المألوف ، وعليه يفضل استخدامه للأسباب المذكورة أعلاه .

شروط تطبيق الاختبار :

1- تتضمن البيانات على عينة عشوائية من المفردات $X_1, X_2, \ldots, X_m$ من المجتمع "1" بدالة توزيع F وعلى عينة عشوائية أخرى $Y_1, Y_2, \ldots, Y_n$ من المجتمع "2" بدالة توزيع G .

2- العينتين مستقلتين عن بعضهما البعض .

3- وحدة القياس على الأقل ترتيبي .

4- إذا وجد اختلاف بين دوال توزيع المجتمعين ، فإن الاختلاف سيكون في موقع التوزيـع ، أى أنه إذا كانت $F(x) \neq G(x)$ فإن $F(x) = G(x+c)$ حيث c مقدار ثابت .

الفرضيات :

إن الفرضيات الآتية ستكون مناسبة عندما يتحقق الشرط الرابع فقط .

أ - اختبار من طرفين :

$H_0 : (E(X) = E(Y)$ أو) المجتمعان لهما توزيعان متطابقان

$H_1 : (E(X) \neq E(Y)$ أو) المجتمعان يختلفان بالنسبة للموقع

ب - اختبار من طرف واحد :

$H_0 : (E(X) = E(Y)$ أو) المجتمعان لهما توزيعان متطابقان

$H_1 : (E(X) < E(Y)$ أو) إن قيم مجتمع X يبدو أنها أصغر من قيم مجتمع Y

جـ - اختبار من طرف واحد :

$H_0 : (E(X) = E(Y)$ أو) المجتمعان لهما توزيعان متطابقان

$H_1 : (E(X) > E(Y)$ أو) إن قيم مجتمع X يبدو أنها أكبر من قيم مجتمع Y

في كثير من التجارب العملية إن الفرق ما بين التوزيعين يعنى أن $P(X < Y)$ لا يساوى $\frac{1}{2}$

وعليه فإنه يمكن صياغة الفرضيات السابقة بدلالـة هذا الاحتمـال إذا لـم يتحقـق الشـرط الرابـع ، وذلك على النحو الآتي :

أ - اختبار من طرفين :

$$H_0 : P(X < Y) = \frac{1}{2}$$

$$H_1 : P(X < Y) \neq \frac{1}{2}$$

ب - اختبار من طرف واحد :

$$H_0 : P(X < Y) \leq \frac{1}{2}$$

$$H_1 : P(X < Y) > \frac{1}{2}$$

جـ - اختبار من طرف واحد :

$$H_0 : P(X < Y) \geq \frac{1}{2}$$
$$H_1 : P(X < Y) < \frac{1}{2}$$

إحصاءة الاختبار :

لحساب إحصاءة الاختبار يتم ضم بيانات العينتين مع بعضهما البعض ثم ترتب مفرداتهما من الأصغر إلي الأكبر وتعطى رتب لهذه المفردات على حسب ترتيبهما في البيانات ، وإذا كانت هناك مفردات متساوية فإنه يعطى متوسط الرتب المعطاة لها كما لو كان لا يوجد تساوي بينهما، وسوف نرمز لرتب مفردات قيم X_i بالرمز $R(X_i)$ حيث $i = 1,2,...,m$ ولرتب مفردات قيم Y_j بالرمز $R(Y_j)$ حيث $j = 1,2,...,n$ و $N = m + n$ ، فإذا كانت معلمة الموقع (المتوسط أو الوسيط) للمجتمع الأول أصغر من معلمة الموقع للمجتمع الثاني ، فإننا نتوقع (إذا كان حجم العينتين متساوي) أن يكون $\sum_{i=1}^{m} R(X_i)$ أصغر من $\sum_{j=1}^{n} R(Y_j)$ ، ولكن إذا كانت معلمة الموقع للمجتمع الأول أكبر من معلمة الموقع للمجتمع الثاني فإننا نتوقع أن يكون $\sum_{i=1}^{m} R(X_i)$ أكبر من $\sum_{j=1}^{n} R(Y_j)$ ، وستكون إحصاءة الاختبار كما يلي :

$$T = \sum_{i=1}^{m} R(X_i) - \frac{m(m+1)}{2}$$

القرار :

إن القرار سوف يعتمد على الفرضية قيد الاختبار وباستخدام جدول (8) بمعلومية m و n و α وذلك على النحو التالي :

1- في الحالة (أ) : نرفض H_0 عندما تكون قيمة T صغيرة جداً أو كبيرة جداً . وعليه نرفض H_0 عندما تكون T أصغر من $\omega_{\frac{\alpha}{2}}$ أو أكبر من $\omega_{1-\frac{\alpha}{2}}$ حيث $\omega_{1-\frac{\alpha}{2}} = mn - \omega_{\frac{\alpha}{2}}$

و $\omega_{\frac{\alpha}{2}}$ تمثل القيمة الجدولية بجدول (8) .

2- في الحالة (ب) : سوف نرفض H_0 عندما تكـون قيمـة T صغيرة جداً . وعليه نرفض H_0 إذا كانت T أصغر من ω_α .

3- في الحالة (جـ) : سوف نرفض H_0 عندما تكون قيمة T كبيرة جداً وعليه نرفض H_0 إذا كانت T أكبر من $\omega_{1-\alpha}$ حيث $\omega_{1-\alpha} = mn - \omega_\alpha$.

مثال (6) : عينة مؤلفة من سبعة عشر طالباً تم اختيارهم بشكل عشوائي للمشاركة في مشـروع بحث علمي ، حيث تم تعليم ثمانية منهم عن طريق الحضـور للدروس بالطريقـة المألوفـة بينما البقية منهم تم تعليمهم عن طريق التعليم الذاتي وذلك من خلال عرض الدروس في شـريط مرئـي مسجل ، وبعد أربعة أسابيع تقدم الطلبة لنفس الاختبار فكانت النتائج كما يلي :

الطريقة المألوفة (X)	64	76	78	94	82	28	82	75	
التعليم الذاتي (Y)	63	77	65	74	48	37	63	95	78

من هذه البيانات هل يمكن القول بوجود فروق معنوية بين درجات المجموعتين ، عند مستوى المعنوية 5 ٪ .

الحـل :

الفرضيات :

$H_0 : (E(X) = E(Y))$ أو درجات مجتمعي المجموعتين متطابقة

$H_1 : (E(X) \neq E(Y))$ أو درجات مجتمعي المجموعتين مختلفة

إحصاءة الاختبار :

لحساب إحصاءة الاختبار ترتب مفردات العينتين من الأصغر إلى الأكبر وتعطى رتب لهذه المفردات على حسب ترتيبهما في البيانات كما يتضح في الجدول التالي :

درجات X	الرتبة $R(X_i)$	درجات Y	الرتبة $R(Y_i)$
28	1		
		37	2
		48	3
		63	4.5
		63	4.5
64	6		
		65	7
		74	8
75	9		
76	10		
		77	11
78	12.5	78	12.5
82	14.5		
82	14.5		
94	16		
		95	17

إذن :

$$\sum_{i=1}^{8} R(X_i) = 1+6+9+10+12.5+14.5+14.5+16 = 83.5$$ وعليه فإن :

$$T = \sum_{i=1}^{8} R(X_i) - \frac{8(9)}{2} = 83.5 - 36 = 47.5$$

القرار :

حيث أن $\alpha = 0.05$ وبالتالي فإنه من جدول (8) وبمعلومية $m = 8$ و $n = 9$ نجد أن $\omega_{0.025} = 16$ ومنها نجد أن $\omega_{0.975} = mn - \omega_{0.025} = 72 - 16 = 56$.

وحيث أن $16 < T < 56$ وعليه لا توجد معلومات كافية لرفض H_0 ، ومستوى المعنوية المشاهد اكبر من 0.20 .

مثال (7): بفرض أن البيانات الآتية تمثل زمن إحتراق نوع معين من المصابيح الكهربائية المنتجة من قبل شركتين مختلفتين (الزمن مقاس بآلاف الساعات).

3.7	2.8	7.1	8.4	6.2	2.7	الشركة X
6.4	6.8	9.1	7.4	6.9	6.8	الشركة Y

من هذه البيانات هل يمكن القول بأن وسيط زمن إحتراق المصـابيح المنتجـة مـن قبـل الشـركة Y أكبر من وسيط الشركة X عند مستوى المعنوية $\alpha = 0.05$.

الحل :

الفرضيات :

بفرض أن $\tilde{\mu}_Y$ ترمـز لوسيط زمـن إحـتراق المصـابيح المنتجـة مـن قبـل الشـركة Y و $\tilde{\mu}_X$ ترمز لوسيط زمن إحتراق المصـابيح المنتجـة مـن قبـل الشـركة X ، إذن الفرضيـة يمكـن كتابتها كما يلي :

$$H_0: \tilde{\mu}_X = \tilde{\mu}_Y$$
$$H_1: \tilde{\mu}_X < \tilde{\mu}_Y$$

إحصاءة الاختبار :

الشركة X	الرتبة $R(X_i)$	الشركة Y	الرتبة $R(Y_i)$
2.7	1		
2.8	2		
3.7	3		
6.2	4		
		6.4	5
		6.8	6.5
		6.8	6.5
		6.9	8
7.1	9		
		7.4	10
8.4	11		
		9.1	12

إذن $\sum_{i=1}^{6} R(X_i) = 1 + 2 + 3 + 4 + 9 + 11 = 30$ وعليه فإن :

$$T=\sum_{i=1}^{6}R(X_i)-\frac{m(m+1)}{2}=30-\frac{6(7)}{2}=30-21=9$$

القرار :

بما أن $\alpha = 0.05$ و $m = 6$ و $n = 6$ ، وعليه فإنه مـن جدول (8) نجد أن $\omega_{0.05}=8$
وبالتالي فإن $T > 8$ وعليه لا نرفض H_0 ، ومستوى المعنوية المشاهد أكبر من 0.05 .

11 - 5 اختبار كروسكل – وليس لتحليل التباين الآحادى بأستخدام الرتب
The kruskal - wallis one - way analysis of variance by ranks

لقد تعرضنا في الفصل العاشر لأسلوب تحليل التبـاين واختبـارات لمقارنـة متوسـطات عدة مجتمعات وذلك على أساس عينات عشوائية تم اختيارها من مجتمعات لها توزيعات طبيعية بتباين مشترك σ^2 . وفي هذا البند سوف نستخدم أسـلوباً لامعلميـاً لمقارنـة عـدة مجتمعـات ولا يتطلـب تطبيقه أي شروط تتعلق بشكل التوزيعات الاحتمالية لهذه المجتمعات .

إن هذا الاختبار يستخدم لاختبار الفرضية التي تهدف إلى معرفة ما إذا كانت عدة عينـات قيد الدراسة قد تم اختيارها (سحبها) من مجتمعـات بـدوال توزيـع متطابقـة ، عـلاوة علـى ذلـك انـه يستخدم أكبر قدر ممكن مـن المعلومـات التـي بالعينـات مقارنـة باختبـارات أخـرى تسـتخدم لنفس الغرض، وعليه فهو أكثر قوة ويفضل استخدامه خاصـة عندمـا تكـون وحـدة قيـاس البيانـات علـى الأقل ترتيبي (ordinal) .

شروط تطبيق الاختبار :

1- تحتوى البيانات علـى مفـردات k عينـة عشـوائية حجـم كـل منهـا $n_1, n_2, n_3, \ldots, n_k$ علـى التوالي ويمكن وضعها في جدول كما يلي :

العينة 1	العينة 2	...	العينة k
x_{11}	x_{21}	...	x_{k1}
x_{12}	x_{22}	...	x_{k2}
$\vdots$	$\vdots$	$\vdots$	$\vdots$
x_{1n_1}	x_{2n_2}	...	x_{kn_k}

2- المفردات مستقلة عن بعضـها البعض خلال العينة ومن عينة إلى أخرى .

3- وحدة القياس على الأقل ترتيبي (ordinal) .

4- إما أن تكون جميع المجتمعات متطابقة التوزيع أو أن بعض المجتمعات تعطى قيم أكبر من المجتمعات الأخرى ، أي أنها مختلفة في الموقع (Location) .

الفرضيات :

دوال التوزيع لجميع المجتمعات (k) متطابقة : H_0

ليس جميع دوال التوزيع متطابقة : H_1

إحصاءة الاختبار:

لحساب إحصاءة الاختبار يتم أولاً استبدال كل مفردة برتبتها وذلك بعد مقارنتها بجميع المفردات من حيث قيمتها العددية في جميع العينات وستعطى الرتبة 1 لأصغر مفردة بجميع العينات ، والرتبة 2 لثاني أصغر مفردة ، ... وهكذا إلى أكبر مفردة حيث يعطى لها الرتبة n ، حيث $n = \sum_{i=1}^{k} n_i$ ، أي أنها تمثل مجموع عدد المفردات في جميع العينات ، وفي حالة وجود مفردات متساوية (متطابقة) سيعطى لها المتوسط الحسابي للرتب المعطاة لها كما لو كان لا يوجد تساوي بينها ، وعليه إذا كان فرض العدم H_0 صحيحاً ، فإننا نتوقع أن يكون توزيع الرتب على العينات محض صدفة وسواء كانت الرتب صغيرة أم كبيرة سوف لن تكون متمركزة في عينة واحدة ، وسيكون أيضاً مجموع الرتب تقريباً متساوي بجميع العينات ، عندما يتم تعديله بالنسبة للعينات التي عدد مفرداتها غير متساوي وستكون إحصاءة الاختبار على النحو الآتي :

$$T = \frac{12}{n(n+1)} \sum_{i=1}^{k} \frac{R_i^2}{n_i} - 3(n+1)$$

حيث $R_i = \sum R(X_{ij})$ = مجموع الرتب المعطاة لمفردات العينة i .

القرار :

عندما يكون عدد العينات ثلاثة فقط وحجم كل منها أصغر من أو يساوى 5 ، فإننا نرفض H_0 إذا كانت قيمة T المحسوبة أكبر من $\omega_{1-\alpha}$ حيث $\omega_{1-\alpha}$ تمثل القيمة الجدولية بجدول (10) ، أما إذا كان عدد العينات أكثر من 3 أو أن عدد المفردات بأحد العينات أكثر من 5 فإننا

رفض H_0 إذا كانت قيمة T المحسوبة أكبر من $\chi^2_{\alpha,k-1}$ حيث $\chi^2_{\alpha,k-1}$ ترمز للقيمة الجدولية بجدول مربع – كاى (6) ، وذلك لأن " كروكسل " برهن على أنه إذا كانت n_i و k كبيرتان فإن الإحصاءة T تتوزع تقريباً وفق توزيع مربع – كاى بدرجات حرية k-1 .الحظ أنه عند وجود تطابق في المفردات ، وكما أشرنا سلفاً يعطى لهم المتوسط الحسابي لرتبتهم وهذا سوف لن يؤثر على قيمة T عندما يكون التطابق خلال العينة نفسها ، ولكن إذا ظهر تطابق بين مفردات العينات المختلفة وكان عدد المفردات المتطابقة صغيراً، فإننا نستخدم نفس الإحصاءة ولكن إذا كان عدد المفردات المتطابقة كبيراً فإنه يفضل استخدام التصحيح للمفردات المتطابقة ولكن سوف لن نتعرض لهذا التصحيح هنا .

المقارنات المتعددة : Multiple comparisons

عند تطبيق الاختبار لمعرفة مدى تطابق عدة مجتمعات وتكون النتيجة هو أن ليست جميع مجتمعات المعاينة متطابقة أي رفض H_0 ، طبيعياً سنتساءل أي المجتمعات تختلف عن بعضها البعض وسيكون المدخل المنطقي للإجابة على مثل هذا السؤال هو استخدام أسلوب أخر مثل اختبار مان – وايتنى ، وذلك لاختبار الفروق المعنوية بين جميع الأزواج الممكنة للعينات ولكن اختبار متوسطات جميع الأزواج الممكنة سيؤثر على احتمال رفض فرض عدم صحيح ، وعليه وجدت طريقة لتطويق هذه المسألة وهى طريقة المقارنات المتعددة ، وهناك عدة طرق تستخدم لهذا الغرض من بينها الطريقة الآتية :

يقال بأن المجتمعين i و j مختلفان إذا تحققت المتباينة الآتية :

$$\left|\frac{R_i}{n_i} - \frac{R_j}{n_j}\right| > z_p \sqrt{\frac{n(n+1)}{12}\left(\frac{1}{n_i} + \frac{1}{n_j}\right)}$$

ويعاد تطبيق هذا الأسلوب مع جميع أزواج المجتمعات ، حيث $p = \dfrac{\alpha}{k(k-1)}$ و z_p ترمز للتجزؤ ذو المرتبة " 1-p " للتوزيع الطبيعي المعياري .

مثال (8) : اختيرت عينات عشوائية من ثلاث أنواع مختلفة من المصابيح الكهربائية وذلك لغرض اختبارها من حيث أيها يعمر أكثر في المتوسط فكانت النتائج كما يلي :

A : 73 , 64 , 67 , 62 , 70

B : 84 , 80 , 81 , 77

C : 82 , 79 , 71 , 75

هل هذه البيانات تشير إلى وجود فروق معنوية بين الأنواع الثلاثة ؟ وإذا وجد فأي الأنواع تختلف عن بعضها البعض ؟

الحـل :

الفرضية :

$$H_0 : \mu_A = \mu_B = \mu_C$$

H_1 : على الأقل نوع واحد مختلف

إحصاءة الاختبار : يتم إعطاء الرتب لمفردات كل عينة بعد مقارنة هذه المفردات مع بعضها وبذلك تكون هذه الرتب كما يلي :

A	B	C
6	13	12
2	10	9
3	11	5
1	8	7
4		
$R_1 = 15$ $n_1 = 5$	$R_2 = 42$ $n_2 = 4$	$R_3 = 33$ $n_3 = 4$

وعليه فإن :

$$T = \frac{12}{n(n+1)} \sum_{i=1}^{3} \frac{R_i^2}{n_i} - 3(n+1)$$

$$= \frac{12}{(13)(14)} (764.45) - 3(14)$$

$$= 8.40$$

القرار :

وحيث أن هناك ثلاث عينات وحجم أكبرها يساوى 5 وعليه نستخدم جدول (10) ، وذلك على حسب ترتيب $n_i : 5 , 4 , 4$. ومن الجدول نجد أن أقرب قيمة لقيمة الاحصاءة T المحسوبة هى $\omega_{0.95} = 5.6176$ ، وحيث أن T أكبر من $\omega_{0.95} = 5.6176$ ، وعليه نرفض H_0 ومستوى المعنوية المشاهد أصغر من 0.01 ، وحيث أنه تم رفض H_0 وعليه نقوم بإجراء المقارنات المتعددة وذلك لتحديد أي الأنواع الثلاث تختلف عن بعضها البعض وذلك كما يلي :

بفرض أن $\alpha = 0.15 \Leftarrow p = \frac{0.10}{3(2)} = \frac{0.10}{6} = 0.025$ ومن جدول التوزيع الطبيعي المعياري نجد أن $z_p = z_{0.025} = 1.96$ وإن

$$z_p \sqrt{\frac{n(n+1)}{12} (\frac{1}{n_i} + \frac{1}{n_j})} = (1.96) \sqrt{\frac{(13)(14)}{12} (\frac{1}{n_i} + \frac{1}{n_j})}$$

وعليه فإن :

المجتمعات	$\left\| \frac{R_i}{n_i} - \frac{R_j}{n_j} \right\|$		$(1.96) \sqrt{\frac{(13)(14)}{12} (\frac{1}{n_i} + \frac{1}{n_j})}$
A و B	7.3	>	5.120
A و C	5.05	<	5.120
B و C	2.25	<	5.397

ومن هذا الجدول يتضح أن A تختلف عن B ولكن A لا تختلف عن C ، و B لا تختلف عن C.

11 - 6 اختبار فريدمان The Friedman test

تعرضنا في البند السابق لأسلوب يمكن تطبيقه على بيانات مأخوذة من ثلاث عينات أو أكثر مستقلة عن بعضها البعض ، وفي هذا البند سنتعرض لأسلوب يستخدم لتحليل بيانات مأخوذة من عدة عينات لها علاقة ببعضها البعض ، وهي عبارة عن تعميم لاختبار في حالة عينتين عندما تكون مرتبطتين مثل اختبار الإشارة . إن مسالة العينات التي لها علاقة ببعضها البعض يمكن وجودها في العديد من التجارب التي تهدف لاكتشاف الفروق ما بين عدة معالجات مختلفة ، حيث يتم تنظيم المفردات في قطاعات (Blocks) تضم الوحدات التجريبية في صورة مجموعات متجانسة مع بعضها البعض ، اي أنه في هذا النوع من التجارب يتم تصنيف وحدات التجربة في قطاعات (Blocks) وبطريقة عشوائية يتم تصنيف المعالجات للقطاعات بحيث كل معالجة يتم تطبيقها مرة واحدة بكل قطاع ، إن مثل هذا الأسلوب يطلق عليه تصميم القطاعات العشوائية الكاملة " Randomized complete block design " . الذي سبق وأن تعرضنا إليه في فصل سابق .وسنعرض الآن أسلوب لامعلمى مناظر لهذا الأسلوب ولكنه لا يشترط أن تكون مجتمعات المعاينة مجتمعات طبيعية علاوة على ذلك قد لا تتوفر لدينا البيانات الأولية للتحليل بل رتب تلك البيانات فقط هذا الاختبار يطلق عليه تسمية "اختبار فريدمان لتحليل التباين ثنائي التصنيف بالرتب " " The friedman two - way analysis of variance by ranks ". وكما يتضح من أسمه فإن هذا الأسلوب يعتمد في حسابه على الرتب المعطاة للمفردات خلال كل قطاع فقط ، وهو سهل التطبيق والفهم .

شروط تطبيق الاختبار :

1- تتضمن البيانات على b من العينات العشوائية المستقلة (قطاعات) حجم كل منها يساوى k ويمكن وضعها في جدول كما يلي:

المعالجات / القطاعات	1	2	3 ⋯ k
1	x_{11}	x_{12}	$x_{13} \cdots x_{1k}$
2	x_{21}	x_{22}	$x_{23} \cdots x_{2k}$
3	x_{31}	x_{32}	$x_{33} \cdots x_{3k}$
⋮	⋮	⋮	⋮ ⋮
b	x_{b1}	x_{b2}	$x_{b3} \cdots x_{bk}$

إن مفهوم المعالجة هنا أكثر عمومية حيث من الممكن أن تعني المعنى المألوف أو أنها تعني مفهوم وضعية معينة مثل الحالة الاقتصادية أو المستوى التعليمي ....الخ .

2- المفردات خلال كل قطاع يمكن ترتيبها بناءً على معيار معين .

3 - لا يوجد تفاعل بين المعالجات والقطاعات .

الفرضيات :

H_0 : جميع المعالجات لها تأثير متطابق (أي أن المجتمعات متطابقة داخل كل قطاع)

H_1 : على الأقل معالجة واحدة يبدو أنها تعطى قيم أكبر من معالجة أخرى

إحصاءة الاختبار :

لحساب إحصاءة الاختبار يتم أولاً استبدال المفردات الأصلية بكل قطاع برتبتها ، أي أنه تتم المقارنة بين المفردات داخل كل قطاع مع بعضها البعض حيث تعطى الرتبة " 1 " لأصغر المفردات بذلك القطاع ، والرتبة " 2 " لثاني أصغر المفردات بنفس القطاع ، وهكذا حتى الرتبة " k " التي تعطى لأكبر المفردات بذلك القطاع ، وبالتالي فإن كل قطاع يتضمن k من الرتب ، ألحظ أن طريقة إعطاء الرتب في اختبار فريدمان تختلف عن الطريقة المتبعة في حالة اختبار كروكسل – وليس الذي يتم فيه إعطاء رتبة للمفردة وذلك من خلال مقارنتها بجميع المفردات

إذا كان H_0 صحيحاً فإننا نتوقع بأن التي بالعينات المختلفة ، وعليه في حالة اختبار فريدمان إذا كان H_1 صحيحاً فإن هذه تكون الرتب موزعة عشوائياً على المعالجات في كل قطاع . أما إذا كان العشوائية في الرتب سوف تنعدم ، أما الخطوة الثانية في حساب الإحصاءة هي إيجاد مجاميع الرتب لكل معالجة بجميع القطاعات ، أي $R_j = \sum_{i=1}^{b} R(X_{ij})$ حيث $j = 1, 2, \ldots, k$ ، وإذا كان H_0 صحيحاً فإننا نتوقع أن تكون هذه المجاميع قريبة من بعضها البعض ، أما إذا كان H_0 خطأ فإننا نلحظ قيمة أحد المجاميع على الأقل ستكون كبيرة مقارنة بالمجاميع الأخرى . وستكون صيغة إحصاءة الاختبار كما يلي :

$$T = \frac{12}{bk(k+1)} \sum_{j=1}^{k} R_j^2 - 3b(k+1)$$

القرار :

سوف نرفض H_0 عند مستوى المعنوية α إذا كانت قيمة T أكبر من أو تساوي $\chi^2_{k-1,\alpha}$ بجدول (6) .

المقارنات المتعددة : Multiple comparisons

عادةً لا يكتفى الباحث بالوقوف عند الخلاصة للقول بأن البيانات تشير لعدم تطابق جميع مجتمعات المعاينة أو أن تأثير المعالجات مختلف ، وإنما يسعى لمعرفة أي المجتمعات (أو المعالجات) تختلف عن بعضها البعض ، وعليه يستخدم طريقة المقارنات المتعددة حيث نقول بأن المعالجتين : i و j مختلفتان إذا تحققت المتباينة الآتية :

$$|R_j - R_i| > z_p \sqrt{\frac{bk(k+1)}{6}}$$

حيث $p = \frac{\alpha}{k(k-1)}$ و z_p ترمز للتجزؤ ذو المرتبة " $1-p$ " للتوزيع الطبيعي المعياري .

مثال (9): أجريت دراسة للمقارنة بين ثلاثة أنواع من الأدوية من حيث سرعة تأثيرها على تنويم المرضى المصابون بأمراض شديدة الألم (الزمن بالدقائق) فكانت النتائج كما يلي :

الدواء \ المريض	1	2	3	4	5	6	7	8	9	10
A	10	10	11	8	7	15	14	10	9	10
B	10	15	15	12	12	10	12	14	9	14
C	15	20	12	10	9	15	18	17	12	16

من هذه البيانات هل يمكن القول بأن للأدوية الثلاثة نفس التأثير عند مستوى المعنوية 5 ٪ .

الحـل :

الفرضية :

H_0 : الأدوية الثلاثة لها نفس التأثير

H_1 : ليس للأدوية الثلاثة نفس التأثير

إحصاءة الاختبار :

فى هذا المثال القطاعات يمثلها الأشخاص لذلك فإن $b = 10$ وعدد المعالجات (الأدوية) تساوى ثلاثة أى أن $k = 3$ ، وعند إستبدال المفردات الأصلية بكل قطاع برتبتها ستكون النتائج كما يلي :

الدواء \ المريض	1	2	3	4	5	6	7	8	9	10	المجموع
A	1.5	1	1	1	1	2.5	2	1	1.5	1	$R_1 = 13.5$
B	1.5	2	3	3	3	1	1	2	1.5	2	$R_2 = 20$
C	3	3	2	2	2	2.5	3	3	3	3	$R_3 = 26.5$

ومن هذا الجدول نجد أن

$$T=\frac{12}{(10)(3)(4)}\left[(13.5)^2+(20)^2+(26.5)^2\right]-(3)(10)(4)$$
$$=\frac{1}{10}\left[182.25+400+702.25\right]-120=8.45$$

القرار :

من جدول مربع - كاى وبدرجات حرية تساوى 2 و $\alpha = 0.05$ نجد أن

$\chi_{0.05,2}=5.991$ ، وحيث أن $T=8.45>5.991$ نرفض H_0 ، أى أن للأدوية الثلاثة تأثير مختلف ، ولتحديد أى الأدوية تختلف عن بعضها البعض سوف نستخدم المقارنات المتعددة وباختيار $\alpha = 0.15$ نجد أن $p=\frac{0.15}{(3)(2)}=0.025$ وعليه فإن $z_p = z_{0.025} = 1.96$ ،

وبالتالى فإن :

$$z_p\sqrt{\frac{bk(k+1)}{6}}=(1.96)\sqrt{\frac{(10)(3)(4)}{6}}=8.765$$

المعالجة	$\lvert R_i - R_j \rvert$	$z_p\sqrt{\frac{bk(k+1)}{6}}$
A و B	6.5	8.765
A و C	13	8.765
B و C	6.5	8.765

وحيث أن A و C أكبر من 8.765 وعليه فإن تأثير A يختلف عن C ولكن تأثير A لا يختلف عن B وتأثير B لا يختلف عن تأثير C .

7 اختبارات جودة المطابقة Goodness - of - fit tests

كثير من المسائل العملية تكون الدراسة متعلقة بطبيعة التوزيع لمجتمع أو أكثر ، وبعبارة تكون الفرضية محل البحث تتعلق بمتغير عشوائي مشاهد له توزيع احتمالي غير معروف، كما نعلم أن صحة أساليب الاستنتاج الإحصائي المعلمي المستخدمة لهذا الغرض تعتمد مثلاً صيغة التوزيع الاحتمالي للمجتمع الذي تم اختيار العينات منه ، وبالتالي عندما تكون هذه غير معروفة وافترضناها كنموذج احتمالي للمجتمع الإحصائي قيد الدراسة يجب التأكد من صحة هذا الافتراض حتى نضمن صحة النتائج التي سوف نصل إليها من جراء تطبيق النموذج ، فمثلاً من الممكن أن تكون الفرضية على النحو التالي : " دالة التوزيع غير معروفة تتبع التوزيع الطبيعي بمتوسط يساوى 5 وتباين يساوى 4 " أو " دالة التوزيع للمتغير X تتبع التوزيع الأسى السالب بمعلمة تساوى 2 " ، إن مثل هذه الفرضيات يمكن باستخدام اختبار جودة المطابقة ، أي اختبار في ما إذا كانت البيانات التي تمثل توزيعاً تجريبياً تتبع توزيعاً احتمالياً معيناً نعتقد بأن هذه البيانات قد تم اختيارها منه ، وفي هذا الدرس المسألة التي يكون فيها محل الاهتمام يتعلق بمعرفة ما إذا كانت بيانات العينة تؤيد القائل بأن مجتمع المعاينة يتبع توزيعاً معيناً .

7-1 اختبار مربع كاى لجودة المطابقة

The chi - square goodness -of - fit test

إن من أشهر وأقدم اختبارات جودة المطابقة هو اختبار مربع كاى لجودة المطابقة ، طرحه بيرسون (1900م) ، وإن هذا الاختبار يشبه اختبارات مربع كاى للاستقلالية التي سبق وأن تعرضنا لها من حيث كون إن إحصاءة الاختبار تنتج من مقارنة التكرارات المتوقعة مع التكرارات المشاهدة ، ولكن اوجه تطبيقها مختلف تماماً .

تطبيق الاختبار :

تمثل البيانات عينة عشوائية بها n من المفردات المستقلة عن بعضها البعض تم اختيارها من مجتمع X ، ويمكن وضع هذه البيانات في جدول توافقي كما يلي

الفئة	1	2	3	.	i	...	r	المجموع
التكرار المشاهد	O_1	O_2	O_3	.	O_i	...	O_r	n

حيث O_i تمثل عدد المفردات التي تقع في الصنف i حيث $i = 1,2,3,...,r$ مـع ملاحظـة أنـه من الممكن أن يكون التصنيف نوعى أو كمي ، فمثلاً من الممكن تصنيف مجموعـة من الأشخاص حسب الجنس (ذكور وإناث) أو من الممكن تصنيفهم حسب العمر (فمثـلاً 20-25 ، 25-30 ، .... الخ) .

2- وحدة القياس على الأقل اسمية (nominal) .

الفرضيات :

إذا رمزنـا لدالـة التوزيـع غير المعروفـة لمجتمـع X بـالرمز $F(x)$ ولدالـة التوزيـع الفرضية بالرمز $F_0(x)$ وهى محددة بالكامل عدا من الممكن أن تكون المعلمـة (أو المعلمـات) غير معروفة ، ويجب تقديرها من بيانات العينة . فإنه يمكن صياغـة الفرضيات الإحصائيـة كما يلي :

$H_0 : F(x) = F_0(x)$ لجميع قيم x

$H_1 : F(x) \neq F_0(x)$ على الأقل لقيمة واحدة من قيم x

أو يمكن إعادة الصياغة كالآتي :

H_0 : العينة تم اختيارها من مجتمع بدالة توزيع $F_0(x)$

H_1 : العينة لم يتم اختيارها من المجتمع الذي دالة توزيعه $F_0(x)$

ألحظ أن الفرض البديل لا يشير إلى كيفيية الاختلاف ما بين دالة التوزيـع الافتراضيـة ودالة التوزيع الحقيقية .

إحصاءة الاختبار :

حيث أنه هناك احتمال بأن تقع أي مفردة يتم اختيارها من المجتمع بأي صنف من التصنيفات المختلفة ، وبالتالي يمكن الرمز لهذه الاحتمـالات بـالرموز $p_1, p_2, p_3, ..., p_r$ على التوالي ، وذلك لأنه يوجد r صنفاً " أو فئة أو خلية " ، وعليه في حالة H_0 يمكن حسـاب التكرار المتوقع بكل صنف كما يلي :

$$E_i = np_i \quad , \quad i = 1,2,3,...,r$$

وبالتالي فإن إحصاءة الاختبار يتم حسابها على النحو التالي :

$$T = \sum_{i=1}^{r} \frac{(O_i - E_i)^2}{E_i}$$

لاحظ أنه عند اختيار العينات من المجتمعات قيد الدراسة فإننا نتوقع بأن تعكس خواص تلك المجتمعات ، وعليه إذا تم اختيار العينة من المجتمع (الفرضي) الذي تم تحديده ، وهذا يعني إن H_0 صحيحاً ، فإننا نتوقع بأن تكون قيم التكرارات المشاهدة (O_i) والتكرارات المتوقعة (E_i) التي تقع في التصنيفات المختلفة قريبة جداً من بعضها البعض ، وخلاف ذلك ستكون الفروق بينهما كبيرة .

القرار :

إذا كانت العينات كبيرة فإن توزيع المتغير T يتبع توزيع مربع-كاى تقريباً بدرجات حرية (r-1) ، وعليه فإننا نرفض H_0 عند مستوى المعنوية α إذا كانت قيمة T المحسوبة أكبر من القيمة $\chi^2_{\alpha,\, r-1}$ حيث $\chi^2_{\alpha,\, r-1}$ تمثل القيمة الجدولية لمربع كاى بجدول (6) ، ألحظ أنه في بعض الأحيان لكي نجرى اختبار مربع كاى لجودة المطابقة يتوجب علينا تقدير معلمات المجتمع الفرضي $F_0(x)$ من بيانات العينة قبل حساب التكرار المتوقع (E_i) ، وعليه فإن درجات الحرية يجب أن يطرح منها عدد المعلمات التي تم تقديرها فإذا كان عدد المعلمات يساوى k مثلاً فإن درجات الحرية في هذه الحالة تساوى (r-k-1) بدلاً من (r-1) .

ملحوظة : كما أشرنا سابقاً أن توزيع T توزيع تقريبي وذلك على افتراض أن حجم العينات كبيراً، فإذا كانت بعض قيم E_i صغيرة فإن توزيع مربع كاى قد يكون غير مناسب في مثل هذه الحالة ولقد أقترح كوكرن (1952) أنه يجب ألا تكون أي قيمة من قيم E_i أصغر من الواحد ولا أكثر من 20 ٪ من قيم E_i أصغر من 5 . وإذا كانت بعض قيم E_i أقل من الواحد فإنه يمكن ضم هذه الصنوف إلى الصنف الذي بجوارها على أن يتم تعديل درجات الحرية تبعاً لذلك .

مثال (10): بفرض أن البيانات الآتية عينة من عدد الحوادث التي حدثت خلال فترة زمنية مدتها 72 ساعة ونتجت عنها أضرار بالغة :

عدد الحوادث	0	1	2	3	4	5	6	7
عدد الساعات	4	10	15	12	12	6	6	7

هل يمكن القول بأن هذه البيانات تتبع توزيع بواسون عند مستوى المعنوية 5 ٪ .

الحـل :

الفرضية : H_0 : البيانات التي اختيرت منها العينة تتبع توزيع بواسون

H_1 : البيانات التي اختيرت منها العينة لا تتبع توزيع بواسون

إحصاءة الاختبار :

نحن نعلم أن توزيع بواسون له الصيغة التالية :

$$p_X(x) = P(X = x) = \begin{cases} \dfrac{e^{-\lambda}\lambda^x}{x!} & , x = 0,1,2,\cdots \\ 0 & ,ow. \end{cases}$$

وحيث أن λ غير معلومة ، وبالتالي يجب تقديرها من بيانات العينة قبل حساب التكرارات المتوقعة ويتم تقدير λ كما يلي :

$$\hat{\lambda} = \frac{0\times4+1\times10+2\times15+3\times12+4\times12+5\times6+6\times6+7\times7}{72} = \frac{239}{72} = 3.32$$

ولإيجاد التكرارات المتوقعة نستخدم دالة مجتمع بواسون وذلك على النحو التالي :

$$p_X(x) = P(X = x) = \begin{cases} \dfrac{e^{-\hat{\lambda}}(\hat{\lambda})^x}{x!} & , x = 0,1,2,3,4,5,6,7 \\ 0 & ,ow. \end{cases}$$

وإن

$$P(X = x+1) = \frac{e^{-\hat{\lambda}}(\hat{\lambda})^x}{x!}.\left(\frac{\hat{\lambda}}{x+1}\right)$$

ومنها نجد أن

$$p_0 = P(X=0) = e^{-3.32} = 0.0362 \;, p_1 = 0.1202 \;,$$
$$p_2 = 0.1995 \;, p_3 = 0.2208 ,\; p_4 = 0.1833 \;,$$
$$p_5 = 0.1217 \;, p_6 = 0.0673 \;, \; p_7 = 0.0319 \;.$$

وعليه فإن التكرار المتوقع $E_i = np_i$ و $i = 0,1,2,3,4,5,6,7$ يكون كالآتي :

$$E_0 = 2.61 \;, E_1 = 8.65 \;, E_2 = 14.36 \;, E_3 = 15.90 \;,$$
$$E_4 = 13.20 \;, E_5 = 8.76 \;, E_6 = 4.85 \;, E_7 = 2.30$$

إن قيمة إحصاءة الاختبار يتم حسابها باستخدام العلاقة الآتية :

$$T = \sum_{i=0}^{7} \frac{(O_i - E_i)^2}{E_i}$$
$$= \frac{(4-2.61)^2}{2.61} + \frac{(10-8.65)^2}{8.65} + \cdots + \frac{(6-4.85)^2}{4.85} + \frac{(7-2.3)^2}{2.3}$$
$$= 0.740 + 0.211 + 0.029 + 0.957 + 0.109 + 0.870 + 0.273 + 9.604$$
$$= 12.793$$

القرار :

من جدول مربع كاى وبدرجات حرية $r - k - 1 = 8-1-1 = 6$ ، ومستوى معنوية 0.05 ، نجد أن $\chi^2_{0.05,6} = 12.592$ ، وحيث أن T أكبر من 12.592 وعليه نرفض H_0 ومستوى المعنوية المشاهد (p value) أصغر من 0.05 .

مثال (11) : بفرض أن البيانات التالية تمثل عينة من قراءات لكميات الأمطار التي سقطت على مناطق مختلفة خلال فترة زمنية معينة :

18.2	21.4	22.6	17.4	17.6	16.7	17.1	21.4	20.1	17.9	16.8	23.1
22.3	21.7	19.6	18.4	17.7	19.3	18.4	18.6	17.8	16.9	21.4	20.6
19.8	18.7	17.5	17.8	18.3	18.9	19.6	20.6	18.7	18.3	18.8	21.4
20.9	21.8	22.6	22.1	21.4	22.3	21.4	23.2	21.6	22.4	19.6	18.6
19.9	20.7	21.8	22.2	21.5	21.1	19.6	18.9	20.8	19.6	20.4	23.0

هل يمكن القول بأن البيانات التي اختيرت منها هذه العينة تتبع التوزيع الطبيعي عندما $\alpha = 0.05$.

الحل :

الفرضية :

H_0 : إن البيانات التي اختيرت منها هذه العينة تتبع التوزيع الطبيعي

H_1 : إن البيانات التي اختيرت منها هذه العينة لا تتبع التوزيع الطبيعي

إحصاءة الاختبار :

حيث أن كلا من μ و σ^2 مجهولتين وعليه يجب تقديرهما من بيانات العينة قبل حساب إحصاءة الاختبار ، ولتقدير هاتين المعلمتين يفضل أن توضع البيانات في جدول تكراري بأطوال فترات متساوية وذلك كما يلي :

الفترات	التكرار (f_i)	مراكز الفترات (x_i)
16.35 - 17.25	4	16.8
17.25 - 18.15	7	17.7
18.15 - 19.05	12	18.6
19.05 - 19.95	8	19.5
19.95 - 20.85	6	20.4
20.85 - 21.75	11	21.3
21.75 - 22.65	9	22.2
22.65 - 23.55	3	23.1

وعليه فإن القيم التقديرية للمعلمتين μ و σ^2 تكونا كما يلي :

$$\hat{\mu} = \frac{\sum_{i=1}^{k} f_i x_i}{n} = \frac{1196.1}{60} \cong 19.94 \qquad , n = \sum_{i=1}^{k} f_i \quad , k = 8$$

$$\hat{\sigma}^2 = \frac{n\sum_{i=1}^{k} f_i x_i^2 - \left(\sum_{i=1}^{k} f_i x_i\right)^2}{n(n-1)} = \frac{(60)(24039.45) - (1196.1)^2}{(60)(59)} = 3.3084$$

و $\hat{\sigma} = 1.82$.

ولحساب p_i نستخدم جدول التوزيع الطبيعي المعياري وذلك على النحو الآتي :

حدود الفترات (b_i)	$z_i = \frac{b_i - 19.94}{1.82}$	$P(Z \le z_i)$	الفترة	p_i	$E_i = np_i$
16.35	-1.97	0.0244	> 16.35	0.0244	$4.16 = \{1.46$
17.25	-1.48	0.0694	16.35 - 17.25	0.0450	2.70
18.15	-0.98	0.1635	17.25 - 18.15	0.0941	5.65
19.05	-0.49	0.3121	18.15 - 19.05	0.1486	8.92
19.95	0.01	0.5040	19.05 - 19.95	0.1919	11.51
20.85	0.50	0.6915	19.95 - 20.85	0.1875	11.25
21.75	0.99	0.8389	20.85 - 21.75	0.1474	8.84
22.65	1.49	0.9319	21.75 - 22.65	0.0930	5.58
23.55	1.98	0.9761	22.65 - 23.55	0.0442	$4.08 = \{2.65$
			23.55 ≥	0.0239	1.43

وعليه يمكن وضع التكرارات المتوقعة والمشاهدة في جدول كما هو مبين أدناه :

الخلية	1	2	3	4	5	6	7	8
التكرار المشاهد	4	7	12	8	6	11	9	3
التكرار المتوقع	4.16	5.65	8.92	11.51	11.25	8.84	5.58	4.08

وبالتالي فإن

$$T = \frac{(4-4.16)^2}{4.16} + \frac{(7-5.65)^2}{5.65} + \frac{(12-8.92)^2}{8.92} + \ldots + \frac{(3-4.08)^2}{4.08} = 7.82$$

القرار :

حيث أنه هناك 8 خلايا " فئات " وتم تقدير معلمتين ، وبالتالي فإن درجات الحرية ستكون (8-2-1=5) ومن جدول مربع كاي وعند مستوى المعنوية $\alpha=0.05$ نجد أن $\chi^2_{0.05,5} = 11.07$ ، وعليه لا توجد معلومات كافية لفرض H_0 ومستوى المعنوية المشاهد (p-value) أكبر من 0.10 ، نود أن نشير هنا إلى أن اختبار مربع كاي لجودة المطابقة قد يكون غير مناسب للتوزيعات المتصلة ولكن يمكن أن نبرر استخدامه في مثل هذه الحالة طالما أن توزيع المجتمع المتصل يمكن وضعه في مجموعة تتضمن عدد محدود من الفترات المنفصلة .

وإن اختبار كولو مجروف - سميرنوف الذي سنتعرض إليه في بند قادم سيكون هو الأنسب للتوزيعات المتصلة .

مثال (12) : ألقيت زهرة نرد 600 مرة فكانت النتائج كما يلي :

الرقم	1	2	3	4	5	6
التكرار	87	96	108	89	122	98

على ضوء هذه البيانات هل يمكن القول بأن زهرة النرد متزنة .

الحـل :

الفرضية :

H_0 : زهرة النرد متزنة

H_1 : زهرة النرد غير متزنة

إحصاءة الاختبار :

إذا كانت زهرة النرد متزنة فإن احتمال (p_i) ظهور أي رقم من الأرقام الستة يساوى $\frac{1}{6}$.

وعليه فإن التكرار المتوقع في هذه الحالة سيكون متساوي في جميع الحالات أي أن :

$$E_i = np_i = (600)\left(\frac{1}{6}\right) = 100 \quad , i = 1,2,3,4,5,6$$

وبالتالي فإن

$$\begin{aligned} T &= \sum_{i=1}^{6} \frac{(O_i - E_i)^2}{E_i} \\ &= \frac{(87-100)^2}{100} + \frac{(96-100)^2}{100} + \ldots + \frac{(98-100)^2}{100} \\ &= 8.58 \end{aligned}$$

القرار :

حيث أنه لم يتم تقدير أي معلمة هنا ، وعليه فإن درجات الحرية تكون 5 =1- 6 ومن جدول مربع-كاى وبدرجات حرية 5 ومستوى معنوية α=0.05 نجد أن $\chi^2_{0.05,5} = 11.07$ وحيث أن T أقل من 11.07 نقبل H_0 أي أن زهرة النرد متزنة ومستوى المعنوية المشاهد (p-value) أكبر من 0.10 .

11 - 7 - 2 اختبار كولو مجروف – سميرنوف لعينة واحدة

Kolmogorov - Smirnov One - Sample Test

إن اختبار مربع كاي لجودة المطابقة الذي تمت مناقشته في البند السابق يستخدم عندما تكون البيانات من نوع أسمى (nominal)، وفى هذا البند سنعرض اختبار يستخدم لاختبار جودة المطابقة في حالة البيانات المتصلة ، وعليه فإنه يمكن استخدامه مع بيانات وحدة قياسها على الأقل ترتيبي (ordinal) . لقد أقترح العالم الروسي كولو مجروف في سنة (1933) اختبار جودة المطابقة في حالة عينة واحدة ، وفى سنة (1939) أقترح العالم الروسي سميرنوف اختبار جودة المطابقة في حالة بيانات تتعلق بعينتين ، وبسبب وجود التشابه ما بين الاختبارين فقد اطلق على الاختبار الأول أسم كولو مجروف – سميرنوف لعينة واحدة ، وعلى الثاني اختبار كولو مجروف –سيمنروف لعينتين .

يعتمد هذا الاختبار على توزعين احتماليين هما التوزيع الاحتمالي التراكمي النظري (المتوقع) والتوزيع الاحتمالي التراكمي التجريبي (المشاهد) ، فعند اختيار عينة عشوائية من مجتمع بتوزيع $F(x)$ غير معروف حيث $F(x) = P(X \leq x)$ فإن الهدف هو تحديد ما إذا كانت $F(x) = F_0(x)$ لجميع قيم x حيث $F_0(x)$ تمثل دالة التوزيع التراكمي الفرضية ، ولتحقيق هذا الهدف فإن اختبار كولو مجروف –سميرنوف لعينة واحدة ينظر إلى التقارب ما بين $F_0(x)$ و $S(x)$ حيث $S(x)$ تمثل دالة التوزيع التراكمي التجريبي ، فإذا كان هذا التقارب ضعيفاً فإنه يعنى عدم صحة الافتراض القائل بأن $F(x) = F_0(x)$ وخلاف ذلك الافتراض صحيحاً .

شروط الاختبار :

تتألف البيانات من عينة عشوائية $X_1, X_2, \cdots, X_n$ عدد مفرداتها يساوى n من مجتمع دالة توزيعه غير معروفة ونرمز لها بالرمز $F(x)$.

الفرضيات :

إذا كانت $F_0(x)$ تمثل دالة التوزيع الفرضية (دالة الاحتمال التراكمي) فإنه يمكن صياغة فرض العدم والبدائل المناظرة كما يلي :

أ – اختبار من طرفين :

$H_0 : F(x) = F_0(x)$ ، لجميع قيم x

$H_1 : F(x) \neq F_0(x)$ ، على الأقل لقيمة واحدة من قيم x

ب - اختبار من طرف واحد :

$H_0 : F(x) \geq F_0(x)$ ، x لجميع قيم

$H_1 : F(x) < F_0(x)$ ، x على الأقل لقيمة واحدة من قيم

جـ - اختبار من طرف واحد :

$H_0 : F(x) \leq F_0(x)$ ، x لجميع قيم

$H_1 : F(x) > F_0(x)$ ، x على الأقل لقيمة واحدة من قيم

إحصاءة الاختبار :

إذا كانت S(x) تمثل دالة التوزيع التجريبي (العيني) ، أى أن S(x) تمثل دالة الاحتمال التراكمي من بيانات العينة ، أي أن

S(x) = نسبة مفردات العينة التى أقل من أو تساوى x .

= (عدد المفردات التى أقل من أو تساوى x) ÷ (n) .

إن إحصاءة الاختبار تعتمد على الفرضية قيد الدراسة وذلك على النحو التالي :

أ - فى حالة اختبار من طرفين : إحصاءة الاختبار تكون كالآتى :

$$T = \sup_x |S(x) - F_0(x)|$$

أى أن T تساوى الحد الأقصى للفرق المطلق ما بين S(x) و $F_0(x)$ لجميع قيم x أى أنه عند تمثيل الدالتين بيانياً فإن T تساوى أقصى مسافة عمودية مطلقة ما بين S(x) و $F_0(x)$.

ب- فى حالة اختبار من طرف واحد وعندما يكون الفرض البديل $H_1 : F(x) < F_0(x)$ فإن إحصاءة الاختبار تكون كالآتى :

$$T^+ = \sup_x [F_0(x) - S(x)]$$

بيانياً ، إن هذه الاحصاءة تمثل الحد الاقصى للبعد ما بين منحني $F_0(x)$ و S(x) عندما يكون منحنى الدالة $F_0(x)$ أعلى من منحنى الدالة S(x) .

جـ - فى حالة اختبار من طرف واحد وعندما يكون الفرض البديل $H_1 : F(x) > F_0(x)$ فإن إحصاءة الاختبار تكون كالآتى :

$$T^{-} = \sup_{x}\left[S(x) - F_0(x)\right]$$

علماً ، إن هذه الاحصاءة تمثل الحد الاقصى للبعد ما بين منحني $F_0(x)$ و $S(x)$ عندما يكون منحنى الدالة $S(x)$ أعلى من منحنى الدالة $F_0(x)$.

ملاحظة :

لحساب T أو T^{+} أو T^{-} إنه من المناسب ترتيب بيانات العينة من الأصغر إلى الأكبر ويرمز لمفردات x المرتبة بالرمز : $x_{(1)} \leq x_{(2)} \leq x_{(3)} \leq \cdots \leq x_{(n)}$ ، فمثلاً إذا كانت $x_1 = 6, x_2 = 5, \quad x_4 = 3, \quad x_3 = 10$ فإن $x_{(1)} = 3, x_{(2)} = 5, \quad x_{(3)} = 6,$ $x_{(4)} = 10,$. والحظ أولاً أن

$$T = \max(T^{+}, T^{-}) \qquad (أ)$$

حيث

$$T^{-} = \max\{\max_{1 \leq i \leq n}[\tfrac{i}{n} - F_0(x_{(i)})], 0\} \qquad (ب)$$

وإن

$$T^{+} = \max\{\max_{1 \leq i \leq n}[F_0(x_{(i)}) - \tfrac{i-1}{n}], 0\} \qquad (جـ)$$

ومن (أ) و (ب) و (جـ) نجد أن

$$T = \max_{1 \leq i \leq n}\{\max[\tfrac{i}{n} - F_0(x_{(i)}), \; F_0(x_{(i)}) - \tfrac{i-1}{n}]\} \qquad (د)$$

حيث $S(x) = \frac{i}{n}$.

القرار :

يرفض H_0 عند مستوى المعنوية α إذا كانت إحصاءة الاختبار T أو T^{+} أو T^{-} أكبر من $\omega_{1-\alpha}$ حيث $\omega_{1-\alpha}$ تمثل القيمة الجدولية (11) الذي يتضمن القيمة الجدولية عند مستوى المعنوية 0.20 ، 0.10 ، 0.05 ، 0.02 ، 0.01 في حالة اختبار من طرفين ومستوى المعنوية 0.10 ، 0.05 ، 0.025 ، 0.01 ، 0.005 في حالة اختبار من طرف واحد فإذا كانت العينة تم اختيارها من المجتمع الفرضي فإن الفرق ما بين القيم المشاهدة لكل من $S(x)$ و $F_0(x)$ سيكون صغيراً وبعبارة أخرى ، إذا كان H_0 صحيحاً فإن الفرق ما بين القيم المشاهدة لكل من $S(x)$

و $F_0(x)$ سيكون صغيراً لجميع قيم x . أما إذا كان H_0 خطأ فإننا نتوقع أن يكون هذا الفرق كبيراً .

ملحوظة : عندما يجب تقدير معالم التوزيع الفرضي من بيانات العينة فإن اختبار كولمجروف - سيمنروف لا يمكن تطبيقه بالمعنى الصحيح وإن النتيجة ستكون تقريبية ، أي أن مستوى المعنوية الحقيقي سيكون أصغر من قيمته الاسمية .

مثال (13) : بفرض أن البيانات الآتية :

0.414	0.523	0.229	0.942	0.097
0.394	0.572	0.486	0.273	0.358

تمثل عينة عشوائية ونود اختبار الفرضية بأن هذه العينة تم اختيارها من التوزيع المنتظم بدالة توزيع تراكمي معرفة كما يلي :

$$F_0(x)=\begin{cases}0 & , x\le 0\\ x & , 0<x<1\\ 1 & , x\ge 1\end{cases}$$

الحـل :

الفرضية :

$H_0 : F(x) = F_0(x)$ ، لجميع قيم x

$H_1 : F(x) \neq F_0(x)$ ، على الأقل لقيمة واحدة من قيم x

إحصاءة الاختبار :

المفردة x_i	$F_0(x_{(i)})$	$\frac{i}{n}$	$\frac{i-1}{n}$	$\frac{i}{n} - F_0(x_{(i)})$	$F_0(x_{(i)}) - \frac{i-1}{n}$
0.097	0.097	0.1	0	0.003	0.097
0.229	0.229	0.2	0.1	0.029	0.129
0.273	0.273	0.3	0.2	0.027	0.073
0.358	0.358	0.4	0.3	0.042	0.058
0.394	0.394	0.5	0.4	0.106	-0.006
0.414	0.414	0.6	0.5	0.186	-0.086
0.486	0.486	0.7	0.6	0.214	-0.114
0.523	0.523	0.8	0.7	0.277	-0.177
0.572	0.572	0.9	0.8	0.328	-0.228
0.942	0.942	1.0	0.9	0.058	-0.042

وعليه من الجدول أعلاه و(4) نجد أن :

$$T = \max_{1 \le i \le 10} \{ \max [\tfrac{i}{n} - F_0(x_{(i)}) \;\; , \; F_0(x_{(i)}) - \tfrac{i-1}{n}] \} = 0.328$$

القرار :

حيث أن 10= n و 0.05 =α من جدول (11) نجد أن $\omega_{0.95} = 0.409$ ، وبما أن $T = 0.328$ أقل من 0.409 وعليه لا توجد معلومات كافية لرفض فرض العدم ، ومستوى المعنوية المشاهد (p-value) أكبر من 0.10 .

ألحظ أنه إذا كانت :

$$H_0 : F(x) \le F_0(x)$$
$$H_1 : F(x) > F_0(x)$$

فإن إحصاءة الاختبار تكون كالآتي :

$$T^- = \max\{ \max_{1 \le i \le 10} [\tfrac{i}{n} - F_0(x_{(i)})] , 0\} = 0.328$$

وحيث أن 10= n و 0.05 =α من جدول (11) نجد أن $\omega_{0.95} = 0.369$ ، وبما أن 0.328 أقل من 0.369 وعليه لا توجد معلومات كافية لرفض فرض العدم .

مثال (14) : بفرض أنه أجريت تجربة لقياس العمر الزمنى لنوع معين من النضائد حيث تم أختيار عينة عشوائية تتكون من 8 نضائد من مجتمع كبير من هذه النضائد فكانت النتائج كمايلى (الزمن مقاس بمئات الساعات) :

3.491 , 1.267 , 2.343 , 0.538 , 5.088 , 5.587 , 2.563 , 3.334

فإذا أفترضنا أن هذه البيانات تشكل عينة عشوائية من توزيع متصل . هل يمكن القول بأنها تتبع التوزيع الاسى بمتوسط يساوى 2 عند مستوى المعنوية 5 ٪ ؟

الحـل :

الفرضية :

$$H_0 : F(x) = F_0(x)$$
$$H_1 : F(x) \neq F_0(x)$$

حيث

$$F(x) = 1 - e^{\frac{-x}{2}}$$

إحصاءة الاختبار :

لحساب هذه الإحصاءة نجرى الحسابات التالية كما هي موضحة في الجدول الأتى :

المفردة x_i	$F_0(x_{(i)})$	$\frac{i}{n}$	$\frac{i-1}{n}$	$\frac{i}{n} - F_0(x_{(i)})$	$F_0(x_{(i)}) - \frac{i-1}{n}$
0.538	0.2359	0.125	0	- 0.1109	0.2354
1.267	0.4693	0.25	0.125	- 0.2193	0.3443
2.343	0.6901	0.375	0.25	- 0.3151	0.4401
2.563	0.7224	0.5	0.375	- 0.2224	0.3474
3.334	0.8112	0.625	0.5	- 0.1862	0.3112
3.491	0.8254	0.75	0.625	- 0.0754	0.2004
5.088	0.9214	0.875	0.75	- 0.0464	0.1714
5.587	0.9388	1.0	0.875	- 0.0612	0.0638

ومن هذا الجدول يتضح أن

$$T^{-} = \max\left\{ \max_{1 \le i \le n} \left[\frac{i}{n} - F_0(x_{(i)}) \right] , \ 0 \right\} = 0.0612$$

و

$$T^{+}=\max\left\{\max_{1\le i\le n}\left[F_0(x_{(i)})-\frac{i-1}{n}\right],0\right\}=0.4401$$

$$\Rightarrow\quad T=\max(T^{-},T^{+})=0.4401$$

القرار :

من جدول (11) و 8 = n و $\alpha = 0.05$ نجد أن $\omega_{0.95}=0.454$ وحيث أن 0.4401 أصغر من 0.454 وعليه لا توجد معلومات كافية لرفض H_0 ، ومستوى المعنوية المشاهد أكبر من 0.05 .

مقارنة بين اختبار مربع-كاى واختبار كولو مجروف – سيمنروف لجودة المطابقة :

في بعض الأحيان يواجه الباحث مسائل يتطلب تحليلها باستخدام أساليب جودة المطابقة ، ولكنه لا يعرف أي من هذه الأساليب يجب أن يستخدم ولكن في معظم الحالات سيكون اختياره إما لاختبار مربع كاي أو لاختبار كولو مجروف – سيمنروف . وبالتالي سوف نعرض بعض أوجه الاختلاف بين هذين الاختبارين في النقاط التالية :

1- إن اختبار مربع كاى يستخدم في حالة البيانات التكرارية ، بينما اختبار كولو مجروف – سيمنروف يستخدم في حالة البيانات المتصلة ، وعند استخدامه في لبيانات منفصلة فإنه سيكون غير دقيق .

2- إن اختبار كولو مجروف -سيمنروف يمكن استخدامه لاختبار فرضيات من طرفين ومن طرف واحد ، بينما اختبار مربع كاى يستخدم في اختبار من طرفين فقط ، وعليه فهو لا يظهر اتجاه وجه الاختلاف ما بين المفردات المشاهدة والمتوقعة .

3- إن التوزيع الاحتمالي لإحصاءة اختبار كولو مجروف – سيمنروف معروف ووضعت له جداول في حالة التوزيعات المتصلة المعرفة تحت فرض العدم (H_0) بينما التوزيع الاحتمالي لإحصاءة مربع كاى تقريبي .

4- إن اختبار مربع كاى يكون مناسباً لبيانات وحدة قياسها أسميه (nominal) وعندما يكون التوزيع الفرضي توزيع منفصل فإنه غالباً ما توجد مثل تلك البيانات في الواقع العملي .

5- يمكن استخدام اختبار كولو مجروف – سيمنروف لإيجاد حدود ثقة لدالة التوزيع $F_{(x)}$.

6- إن اختبار مربع كاي يتطلب أن توضع البيانات في مجموعة من الفئات " التصنيفات " لها علاقة بالمتغير محل البحث ، بينما اختبار كولو مجروف - سيمنروف لا يتطلب ذلك وعليه فهو يستخدم البيانات بشكل أكثر كفاءة من اختبار مربع كاي .

11- 8 اختبار العشوائية لعينة واحدة The one - sample runs test for randomness

إن معظم الأساليب الإحصائية التي تستخدم لدراسة ظاهرة معينة بمجتمع ما ، تفترض أن مفردات العينة التي يتم اختيارها من ذلك المجتمع ستكون عشوائية حتى يكون للاستنتاج الإحصائي معنى ، فإذا وجد شك في عدم صحة هذا الافتراض يجب أن يكون لدينا أسلوب علمي للتحقق من ذلك ، فمثلاً عند استخدام أساليب مراقبة الجودة نقوم برسم خرائط للتحكم ودراسة عدد الوحدات المعيبة في الإنتاج ، وللقيام بذلك عادةً ما تؤخذ عينات من الإنتاج بشكل دوري ومعرفة عدد الوحدات المعيبة ، وبالتالي قد يكون هذا العدد أكبر أو أصغر من العدد المسموح به وإن الهدف من وراء ذلك هو معرفة ما إذا كان هذا العدد من الوحدات المعيبة بالإنتاج التي ظهرت بالعينات المختارة يحدث بشكل عشوائي لأنه إذا لم يكن عشوائي فهو مؤشر على ضعف التحكم في الإنتاج ، ولقد أقترح أسلوب لاختبار العشوائية في مثل هذه الحالة يطلق عليه أسم اختبار الدورات للعشوائية بعينة واحدة " The one - sample runs test for randomness " .

إن الأساليب أو الطرائق التي تستخدم لدراسة العشوائية بظاهرة معينة تعتمد أساساً على طبيعة وعدد الدورات الموجودة في بيانات (أو مفردات) تلك الظاهرة ، وتعرف الدورة على أنها متتابعة من العناصر المتشابهة تسبق وتلحق بعناصر من نوع آخر وعدد العناصر داخل كل الدورة يطلق عليها طول الدورة . وغالباً ما نشك في عشوائية السلسلة التي تمثل ظاهرة معينة وخاصة إذا كان عدد الدورات قليلاً أو كثيراً ، فمثلاً عند إلقاء عملة متزنة ثمانية مرات من الممكن مشاهدة النتائج الآتية : HHHHTTTT أو من الممكن مشاهدة الآتي : HTHTHTHT

ففي الحالتين نشك في العشوائية وذلك لأنه في الحالة الأولى توجد دورتين فقط هما HHHH و TTTT بينما في الحالة الثانية توجد ثمانية دورات ويجب أن نشير هنا إلى أننا سوف نقتصر على دراسة العشوائية بالتجارب التي يكون بها نوعين من النتائج فقط أما إذا كانت النتائج يمكن تصنيفها إلى أكثر من نوعين فإن الاختبار الذي سنتناوله هنا سيكون غير مناسب .

شروط تطبيق الاختبار :

تتضمن البيانات متتابعة من المفردات مرتبة على حسب حدوثها ، ويمكن تصنيفها إلى نوعين منفصلين فقط .ولنفرض أن n تمثل حجم العينة و n_1 تمثل عدد مفردات أو مشاهدات النوع الأول و n_2 تمثل عدد مفردات أو مشاهدات النوع الثاني .

الفرضيات :

ا - اختبار من طرفين :

H_0 : نتائج حدوث النوع الاول والثاني من المفردات عشوائية

H_1 : نتائج حدوث النوعين غير عشوائية

ب - اختبار من طرف واحد :

H_0 : نتائج حدوث النوع الاول والثاني من المفردات عشوائية

H_1 : نتائج حدوث النوعين غير عشوائية

(وذلك لأنه هناك عدد قليل من الدورات التى يمكن إيعازها للصدفة) .

جـ - اختبار من طرف واحد :

H_0 : نتائج حدوث النوع الاول والثاني من المفردات عشوائية

H_1 : نتائج حدوث النوعين غير عشوائية

(وذلك لأنه هناك عدد كبير من الدورات التى يمكن إيعازها للصدفة) .

إحصاءة الاختبار :

إحصاءة الاختبار تساوى عدد الدورات أى أن T = عدد الدورات .

القرار :

1- فى الحالة (أ) : بمعلومية n_1 و n_2 وعند مستوى المعنوية 5٪ ومن جدول (a -14) نوجد القيمة الحرجة الصغرى (t_1) ومن جدول (14-b) نوجد القيمة الحرجة العليا (t_2) ونرفض H_0 إذا كانت T أصغر من أو تساوى (t_1) او اكبر من أو تساوى (t_2) .

2- فى الحالة (ب) : بمعلومية n_1 و n_2 وعند مستوى المعنويـة 0.025 ومن جدول (14-a)
نوجد القيمة الحرجة t ونرفض H_0 إذا كانت T أصغر من أو تساوى t .

3- فى الحالة (جـ) : بمعلومية n_1 و n_2 وعند مستوى المعنوية 0.025 ومن جدول (14-b)
نوجد القيمة الحرجة t ونرفض H_0 إذا كانت T أكبر من أو تساوى t .
إذا كانت قيمة n_1 أو n_2 ليست بجدول (14-a) أو جدول (14-b) نستخدم أقرب قيمة لها .

ملحوظة :
إذا كانت n_1 أو n_2 أكبر من 20 فإننا نستخدم التقريب لإيجاد القيمة الحرجة وذلك على النحو الآتى :

بتعريف : $Z = \frac{T - \mu}{\sigma}$

حيث : $\mu = \frac{2n_1 n_2}{n_1 + n_2} + 1$ و $\sigma = \sqrt{\frac{2n_1 n_2 (2n_1 n_2 - n_1 - n_2)}{(n_1 + n_2)^2 (n_1 + n_2 - 1)}}$

والمتغير Z يؤول الى التوزيع الطبيعى المعيارى عندما يكون H_0 صحيحاً ثم نقارن قيمة Z مع القيمة الجدولية للتوزيع الطبيعى المعيارى عند مستوى المعنوية المطلوب .

مثال (15) : فى عينة تتكون من 16 وحدة من وحدات إنتاج أحد الآلات بمصنع ما وجد أن بها وحدات غير تالفة (N) ووحدات تالفة (D) وذلك على النحو الآتى :

NNDDNNNNDDNNNDDD

ونود معرفة ما إذا كان إنتاج هذه الآلة يظهر بشكل عشوائى عند مستوى المعنوية 5٪ .

الحـل :

إن هذه البيانات تفى بشروط تطبيق هذا الاختبار وذلك لأنها سجلت كما تمت مشاهدتها أثناء الانتاج ولنفرض أن n_1 = عدد الوحدات غير التالفة (9) و n_2 = عدد الوحدات التالفة (7) وعليه فإن :

الفرضية :

H_0 : إنتاج الوحدات غير التالفة والتالفة يظهر بشكل عشوائى

H_1 : الإنتاج غير عشوائى

إحصاءة الاختبار :

من هذه البيانات نجد أن هناك 6 دورات وذلك كما يلي :

$$\underset{1}{\underline{NN}} \quad \underset{2}{\underline{DD}} \quad \underset{3}{\underline{NNNN}} \quad \underset{4}{\underline{DD}} \quad \underset{5}{\underline{NNN}} \quad \underset{6}{\underline{DDD}}$$

وعليه فإن : $T = 6$

القرار :

من جدول (14-a) و (14-b) وبمعلومية $n_1 = 9$ و $n_2 = 7$ نجد أن $t_1 = 4$ و $t_2 = 14$ وحيث أن T تقع ما بين t_1 و t_2 وعليه لا توجد معلومات كافية لرفض H_0 اى أن الانتاج عشوائى .

مثال (16) : في دراسة عن تدفق الدم في شعيرات الأعصاب بالرئة لعينة تتكون من 16 مريض يعانون من ضعف فى الأعصاب والعضلات . سجلت النتائج على حسب جنس المريض، ذكر (M) أم أنثى (F) وذلك كما يلي : FFFMFFMMMFFFFFFM

أختبر الفرضية أن هذه المتتابعة عشوائية عندما $\alpha=0.05$.

الحل :

الفرضية :

H_0 : نتيجة هذه الدراسة عشوائية

H_1 : نتيجة هذه الدراسة غير عشوائية

إحصاءة الاختبار :

من هذه البيانات نلاحظ أن

$$\underset{1}{\underline{FFF}} \quad \underset{2}{\underline{M}} \quad \underset{3}{\underline{FF}} \quad \underset{4}{\underline{MMM}} \quad \underset{5}{\underline{FFFFFF}} \quad \underset{6}{\underline{M}}$$

وعليه توجد 6 دورات وبالتالى فإن $T = 6$

وإن n_1 (عدد الإناث) =11 و n_2 (عدد الذكور) =5

القرار :

من جدول (14-a) و (14-b) وبمعلومية n_1 و n_2 نجد أن $t_1 = 5$ و $t_2 = 11$ وحيث أن T تقع ما بين t_1 و t_2 وبالتالي لا توجد معلومات كافيه لرفض H_0 عند مستوى المعنوية 5 % ، أى أن نتيجة هذه الدراسة عشوائية .

11 - 9 معامل سبيرمان لإرتباط الرتب : The Spearman Rank Correlation Coefficient

لقد تعرضنا في فصل سابق للتوزيع الأحتمالى المشترك للمتغيرين العشـوائين X و Y ، ولقد أوضحنا أن هناك بعض الحالات التى توجد فيها علاقة سببية ما بينهما ، فمثلاً العلاقة ما بين المحصول الزراعي وكمية الأمطار ، ولكن في بعض الأحيان لا توجد علاقة سببية ومع ذلك يوجد ارتباط ما بين المتغيرين . إن المقياس المألوف لقياس هذه العلاقة هو معامل الارتباط ρ الذى سبق وأن تعرضنا إليه فى فصل سابق ، فإذا كان التوزيع المشترك للمتغيرين X و Y معروف فإنه يمكن حساب ρ الذى يقيس قوة العلاقة الخطية ما بين هذين المتغيرين . وإنه من الممكن أن تكون $\rho = 0$ بالرغم من وجود علاقة غير خطية ، ولكن دائماً صحيح القول أنه إذا كان X و Y مستقلين فإن $\rho = 0$.

إن معامل إرتباط العينة (r) الذى سبق وأن تعرضنا إليه فى فصل سابق يمكن استخدامه كتقدير بقيمة واحدة لمعامل الارتباط ρ ، ويمكن إيجاد فترة ثقة حول ρ ، إذا تم الحصول على بيانات من توزيع طبيعي ثنائي ، إن لهذا التوزيع خاصية وهى أنه $\rho = 0$ إذا وإذا فقط كان X و Y مستقلين ، وبالتالى فإن اختبار $\rho = 0$ متطابق لاختبار الاستقلالية ما بين المتغيرين ، ولكن إن شرط أن تكون البيانات من توزيع طبيعي ثنائي لا يمكن تبريره فى كثير من الأحيان .وفى هذا البند سوف نتعرض لطريقة كيفية قياس أو اختبار وجود علاقة بين متغيرين عندما لا يتحقق شرط التوزيع الطبيعي الثنائي .

إن مقياس العلاقة الذي يمكن استخدامه لاختبار $\rho = 0$ عندما لا يتحقق شرط أن تكون البيانات من توزيع طبيعي ثنائي هو معامل سبيرمان للارتباط r_s ، إن هذا المعامل يتم الحصول عليه باستخدام معامل ارتباط العينة مستخدمين رتب القيم المشاهدة بدلاً من القيم نفسها ، فإذا كانت $R(X_i)$ تمثل رتبة X_i فى العينة $X_1, X_2, X_3, \dots, X_n$ وكانت $R(Y_i)$ تمثل رتبة Y_i فى العينة $Y_1, Y_2, Y_3, \dots, Y_n$ فإن معامل سبيرمان لارتباط الرتب معرف كما يلي :

$$r_s = 1 - \frac{\sum_{i=1}^{n} d_i^2}{n(n^2 - 1)} \tag{1}$$

حيث

$$d_i = R(X_i) - R(Y_i)$$

إن معامل الارتباط " r_s " ومعامل الارتباط العادي " r " لهما خواص كثيرة مشتركة ، ولكن في بعض الأحيان من الممكن أن تكون رتب X ورتب Y فى علاقة تامة أي أن $r_s = 1$ ولكن

البيانات الأصلية غير مرتبطة خطياً أى أن $r \neq 1$ ومع ذلك r_s و r كلاهما مقياس لقوة العلاقة ويفيان بالشروط المطلوب توفرها فى أى مقياس يقيس الارتباط بين ظاهرتين .

الحظ أنه إذا وجدت مفردات متساوية داخل العينة الواحدة وكان عددها كبيراً فإننا نستخدم التصحيح التالي للمفردات المتساوية :

$$T = \frac{t^3 - t}{12}$$

حيث t = عدد المفردات المتساوية لرتب معينة ، وإن

$$r_s = \frac{R_1 + R_2 - \sum_{i=1}^{n} d_i^2}{2\sqrt{R_1 . R_2}} \qquad (2)$$

حيث

$$R_1 = \frac{n^3 - n}{12} - \sum T_x$$

و

$$R_2 = \frac{n^3 - n}{12} - \sum T_y$$

حيث T_x = مجموع قيم T للرتب المتساوية والمختلفة فى قيم X .

و T_y = مجموع قيم T للرتب المتساوية والمختلفة فى قيم Y .

الحظ أنه إذا لم يكن عدد المفردات المتساوية كبيراً فإن وجه الاختلاف ما بين استخدام الصيغة (1) أو (2) صغير جداً .

مثال (17) : باحث بقسم علم الحيوان قام بدراسة لتحديد نوعية العلاقة ما بين عدد الأرانب (X) التي في الحضانة الواحدة ، ومتوسط أوزان مواليدها (Y) حيث تم اختيار عينة تتكون من 20 دار حضانة لنوع معين من الأرانب فكانت النتائج كما يلي :

رقم الحضانة	X	Y	رقم الحضانة	X	Y
1	2	47.0	11	4	47.7
2	3	42.3	12	3	47.7
3	6	39.7	13	6	37.1
4	5	40.3	14	1	52.7
5	4	41.4	15	2	42.3
6	7	39.4	16	1	59.2
7	8	40.5	17	6	48.8
8	1	54.2	18	3	51.2
9	2	49.4	19	2	51.0
10	3	43.1	20	3	49.0

على ضوء هذه البيانات هل يمكن القول بوجود علاقة ما بين عدد الأرانب في الحضانة ومتوسط أوزن مواليدها ؟

الحل :

حيث أن عدد الأرانب بكل حضانة صغير ، وبالتالى فإن هذه القيم لا يمكن أن تكون من مجتمع طبيعي ، وعليه فإن X و Y معاً لا يمكن أن يكونا توزيع طبيعي ثنائي ، وبالتالي لحساب قيمة r_s ، يجب إيجاد رتب قيم X ($R(X_i)$) ، ورتب قيم Y ($R(Y_i)$) مع إعطاء متوسط الرتب للقيم المتساوية كما هو مبين في الجدول الآتي :

رقم الحضانة	$R(X_i)$	$R(Y_i)$	رقم الحضانة	$R(X_i)$	$R(Y_i)$
1	5.5	11	11	12.5	10
2	9.5	7.5	12	9.5	12
3	17	3	13	17	1
4	14.5	4	14	2	18
5	12.5	6	15	5.5	7.5
6	19	2	16	2	20
7	20	5	17	17	13
8	2	19	18	9.5	17
9	5.5	15	19	5.5	16
10	14.5	9	20	9.5	14

وعليه فإن :

$d_i = R(X_i) - R(Y_i)$	d_i^2	$d_i = R(X_i) - R(Y_i)$	d_i^2
-5.5	30.25	2.5	6.25
2.0	4.0	-2.5	6.25
14.0	196	16	256
10.5	110.25	-16	256
6.5	42.25	-2.0	4
17	289	-18	324
15	225	4.0	16
-17	289	-7.5	56.25
-9.5	90.25	-10.5	110.25
-4.5	20.25	-4.5	20.25

ومن هذا الجدول نجد أن : $\sum_{i=1}^{20} d_i^2 = 23515$

إذن قيمة r_s باستخدام الصيغة (1) ستكون كالآتي :

$$n = 20 \Rightarrow n^2 = (20)^2 = 400 \Rightarrow n^2 - 1 = 399$$

$$\Rightarrow r_s = 1 - \frac{6 \times 2351.5}{(20)(399)} = 1 - 1.768 = -0.768$$

وإن قيمة r_s باستخدام الصيغة (2) ستكون كالآتي :

X: (5.5)4 , (9.5)4 , (14.5)2 , (12.5)2 , (17)3 , (2)3

Y: (7.5)2

$$T_x = \frac{4^3 - 4}{12} = 5 \quad ,T_x = \frac{3^3 - 3}{12} = 2 \quad ,T_x = \frac{2^3 - 2}{12} = 0.5$$

$$\Rightarrow R_1 = \frac{n^3 - n}{12} - \sum T_x = \frac{(20)^3 - 20}{12} - (2 \times 5 + 2 \times 2 + 2 \times 5) = 649.0$$

$$T_y = \frac{2^3 - 2}{12} = 0.5 \quad ,R_2 \frac{n^3 - n}{12} - \sum T_y = \frac{(20)^3 - 20}{12} - 0.5 = 664.5$$

وعليه فإن :

$$r_s = \frac{649.0 + 664.5 - 2351.5}{2\sqrt{(649.0)(664.5)}} = \frac{-1038}{1313.409} = -0.790$$

الحظ أن هناك فرق بين قيمتى r_s باستخدام الصيغتين وذلك لأن عدد المفردات المتساوية فى X كبيراً .

كما أشرنا سابقاً شرط أن تكون البيانات من توزيع طبيعي ثنائي قد لا يتحقق فى كثيراً من الأحيان ، وبالتالى لاختبار استقلالية المتغيرين العشوائيين X و Y يمكن استخدام معامل سبيرمان لارتباط الرتب وذلك كما يلى :

فإذا كانت (X_i, Y_i) و $i = 1, 2, \cdots, n$ تشكل عينة عشوائية ثنائية من توزيع مشترك متصل (إن شرط أن تكون المعاينة من توزيع متصل وذلك لتفادى مسألة تساوى المفردات) ، وكان المقياس على الأقل ترتيبي ، وأيضاً كان فرض العدم الذى ينص على استقلالية المتغيرين صحيحاً ، فإن توزيع r_s يمكن إيجاده ، وعليه يمكن استخدام r_s كإحصاءة لاختبار ما إذا كان المتغيرين مستقلين أم لا . فإذا كانت ρ_s ترمز لمعامل سبيرمان لارتباط الرتب للمجتمع الإحصائي قيد البحث فإنه يمكن اختبار الفرضيات الآتية :

1- $$H_0 : \rho_s = 0$$
$$H_1 : \rho_s \neq 0$$

إن هذه الفرضية تستخدم إذا كان الهدف هو معرفة ما إذا كان هناك شك فى عدم الإستقلالية .

2- $$H_0 : \rho_s = 0$$
$$H_1 : \rho_s > 0$$

هذه الفرضية تهدف الى معرفة ما إذا كانت العلاقة طردية أم لا .

$$H_0 : \rho_s = 0$$

-3 $$H_1 : \rho_s < 0$$

إن هذه الفرضية تهدف الى معرفة ما إذا كانت العلاقة عكسية أم لا .

إحصاءة الاختبار :

أ - إذا كان حجم العينة (n) ما بين 4 و30 فإن إحصاءة الإختبار هي الصيغة (1) أو (2) وذلك على حسب ما إذا كان عدد المفردات المتساوية قليلة أو كثيرة .

ومن جدول (12) : نرفض H_0 في الحالة (1) إذا كانت r_s أكبر من $r_{s,\frac{\alpha}{2}}$ أو أقل من $-r_{s,\frac{\alpha}{2}}$.

ونرفض H_0 في الحالة (2) إذا كانت $r_s > r_{s,\alpha}$ وفي الحالة (3) نرفض H_0 إذا كانت $-r_s < r_{s,\alpha}$.

ب - إذا كان حجم العينة (n) أكبر من 30 فإن إحصاءة الاختبار تكون كالآتي :

$$Z = r_s \sqrt{n-1}$$

حيث Z تتوزع تقريباً وفق التوزيع الطبيعي المعياري ، وكالعادة نستخدم جداول التوزيع الطبيعي المعياري لتحديد المنطقة الحرجة .

مثال (18) : من بيانات المثال السابق هل يمكن القول بوجود علاقة عكسية ما بين عدد الأرانب بالحضانة ومتوسط وزن المواليد عند مستوى المعنوية 5 ٪ ؟

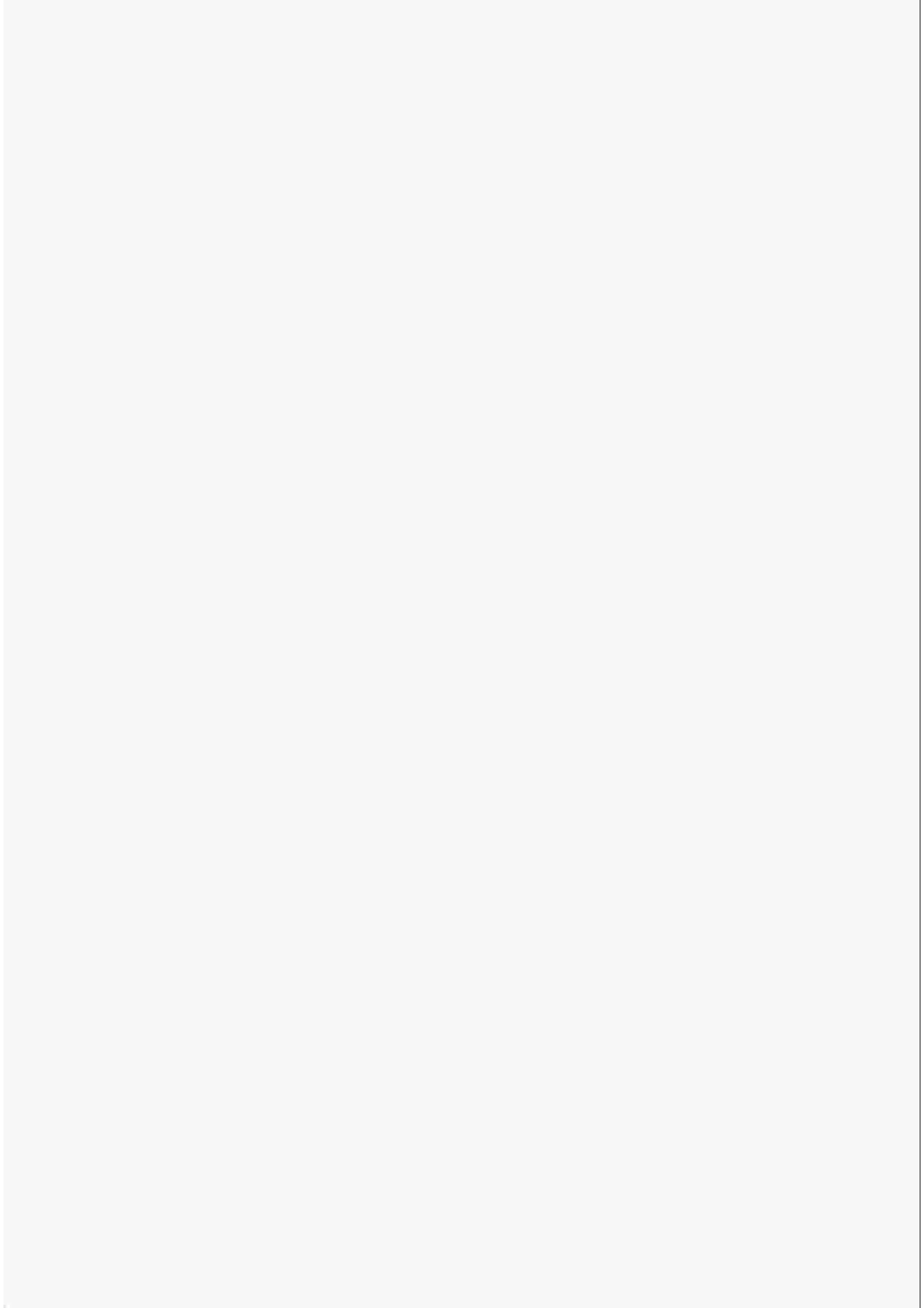
الحـــل :

لاختبار الفرضية أن عدد الأرانب ومتوسط وزن المواليد مستقلين مقابل الفرض البديل بوجود علاقة عكسية بينهما نتبع الآتي :

$$H_0 : \rho_s = 0$$
$$H_1 : \rho_s < 0$$

وعليه إذا كانت α=0.05 فمن جدول (12) نجد أن $-r_{s,0.05,20} = -0.3789$

وحيث أن $-0.3789 > -0.768$ وعليه نرفض H_0 ويمكن القول من المحتمل وجود علاقة عكسية ما بين المتغيرين .الحظ أنه إذا استخدمنا الناتج من الصيغة (2) فإن الخلاصة واحدة .



تمـــرينات Exersises

1 - بفرض أن البيانات التالية تمثل متوسط معدلات الزيادة في الأنتاجية بمجـال الصناعـة خـلال 18 سنة : 0.2 ، 2.5 ، 1.8 ، 1.5 ، 2.8 ، 2.1 ، 2.6 ، 7.1 ، 2.4 ، 2.3 ، 2.4 ، 0.5 ، 2.1 ، 2.6 ، 2.1 ، 2.7 ، 2.0 ، 0.4 من هذه البيانات هل يمكن القول بـأن وسـيط متوسـط معدلات الزيادة في الأنتاجية قد زاد عن 2.0 عندما $\alpha = 0.05$ ؟ استخدم اختبار الاشارة .

2 - يعتقد أستاذ فـي علـم النفس أن ترتيـب الأسئلة لـه تـأثير علـى فرصـة الإجابـة الصحيحـة . ولتأكيد اعتقاده قام بتقسيم 13 طالباً بطريقة عشوائية إلى مجموعتين A و B ، وكان عدد طلبـة المجموعة A " 7 " وعدد طلبة المجموعـة B " 6 " وكـانت أسـئلة المجموعـة الأولـى مرتبـة من الأسهل إلى الأصعب والمجموعة الثانية بالعكس فكانت درجات المجموعتين كما يلى :

Test A : 90 ، 71 ، 83 ، 82 ، 75 ، 91 ، 63

Test B : 66 ، 78 ، 50 ، 68 ، 80 ، 60

من هذه البيانـات هـل يمكـن القـول بـأن درجـات طلبـة المجموعـة A أفضـل مـن درجـات طلبـة المجموعة B عندما $\alpha = 0.05$ ؟ استخدم اختبار مان-وايتني .

3 - البيانات التالية تم جمعها عن 12 حادث صناعى لفترة أسبوع قبل وبعد عمل برنامج تثقيفى مكثف حول الأحتياطيات الواجب مراعاتها للمحافظة على سلامة العاملين بالمصنع :

المصنع	1	2	3	4	5	6	7	8	9	10	11	12
قبل	3	4	6	3	4	5	5	3	2	4	4	5
بعد	2	1	3	5	4	2	3	3	0	3	1	2

من هذه البيانات هل يمكن القول بأن للبرنامج التثقيفى دور في الحد من عدد الحوادث بالمصانع عندما $\alpha = 0.05$ ؟ استخدم اختبار الاشارة .

4 - إن المؤشرات الاقتصادية تعطى قياسات للتغير في الاقتصـاد فـإذا كـانت البيانـات التاليـة تم الحصـول عليهـا فـي الأسبوع الأول مـن ينـاير 1995 والأسبوع الأول مـن شـهر ينـاير 1996 وبمقارنة مجموعة المؤشرات يمكن الحصـول علـى معلومـات تتعلق بـالتغيرات التـي حدثـت فـي الاقتصاد خلال عام 1995 .

السلعة	الرسوم الجمركية	الطاقة الكهربائية	الخشب	النفط الخام	السيارات	الحديد
يناير 1995	190.8	50.8	91.4	98.3	66.4	47.2
يناير 1996	217.1	57.2	59.7	97.8	107.8	66.3

من هذه البيانات هل يمكن القول بأن هناك تغير في المؤشرات الاقتصادية عندما $\alpha = 0.05$ ؟ استخدم اختبار ولكاكسن .

5 - تضمن أحد الاختبارات لمادة الإحصاء التطبيقي 50 سؤالاً ، وكل سؤال له إجابتان إحداهما صحيحة والأخرى خاطئة وكان نموذج الإجابة لهذا الاختبار كما يلي :

F F T T F T T T T T T F T F T T F F F F T T F T T T T T F F F T T F F T T T T
F F T F T F F F F T F

حيث F تعني خطأ و T تعني صحيح . من هذه البيانات هل يمكن القول بأن الأستاذ وضع الإجابة بطريقة عشوائية ؟ عندما $\alpha = 0.05$.

6 - مهندس كهربائي يرغب في مقارنة العمر الزمني لأربعة أنواع من المكثفات ، وللقيام بذلك أختار عينة عشوائية تتضمن خمسة مكثفات من كل نوع ووضعت في الاختبار وتحت نفس الظروف فكانت رتب العمر الزمني المشاهد لكل نوع كما يلي :

النـــــوع

1	2	3	4
1	12	8	14
5	2	9	15
6	17	3	16
7	19	11	4
10	20	13	18

من هذه البيانات هل يمكن القول بأن أحد الأنواع عمره الزمني أكبر من الأنواع الأخرى عندما $\alpha = 0.01$ ؟ (حيث أن بعض الأعمار الزمنية ستكون كبيرة جداً مقارنة بالأخرى وبالتالي فإن التوزيع الطبيعي سوف لن يكون نموذجاً مناسباً هنا) استخدم اختبار كروسكل .

7 - تستخدم إحدى الشركات نوع من الآلات لإنتاج سلعة معينة ، وحيث أنه توفر نوع جديد من الآلات التي يمكن استخدامها لإنتاج نفس السلعة قررت إدارة الشركة شراء هذا النوع ، ولكن قبل أن تشتريه ترغب في مقارنة النوعين القديم (A) والجديد (B) من حيث أيهما أكثر كفاءة ، فقامت باختبار مجموعة من الآلات الجديدة وسجلت عدد مرات توقف النوعين عن العمل في كل أسبوع ولمدة سبعة أسابيع فكانت النتائج كالتالي (علماً بأن الشركة استخدمت نفس العدد من الآلات من النوعين) :

نوع الآلة	الأسبوع 1	2	3	4	5	6	7
A	14	17	10	15	14	9	12
B	12	13	14	12	9	11	11

من هذه البيانات هل يمكن القول بأن هناك اختلاف ما بين النوعين من حيث عدد مرات التوقف عن العمل عندما $\alpha = 0.05$ ؟ استخدم اختبار مان-وايتني .

8 - قام مدير إدارة المراقبة على الجودة في أحد المصانع باختيار عينات عشوائية من إنتاج ثلاثة خطوط خلال عشرة ساعات وسجل عدد القطع التي بها خلل بكل عينة فكانت كما يلي :

خط الإنتاج

1	2	3
6	34	13
38	28	35
3	42	19
17	13	4
11	40	29
30	31	0
15	9	7
16	32	33
25	39	18
5	27	24

من هذه البيانات هل يمكن القول بأن هناك فرق ما بين الخطوط الثلاثة من حيث القطع التي بها خلل عندما $\alpha = 0.05$ ؟ استخدم اختبار كروسكل .

9 - تعتزم الشركة العامة للأطارات أستيراد نوع معين من الأطارات الذي يساهم في أقتصاد البنزين ولتقرير ذلك قام بتجريب نوعين من الأطارات A و B حيث قامت بوضع النوع A في 12 سيارة وتمت قيادتها من قبل سائقين لم يتم تغيرهم حتى أثناء مدة الأختبار ثم قامت بتغيير الأطارات حيث وضعت النوع B وتمت قيادة السيارات من قبل نفس السائقين حتى أنتهت مدة الأختبار فكان أستهلاك البنزين لكل كيلومتر كما يلي :

السيارات	1	2	3	4	5	6	7	8	9	10	11	12
الأطار A	4.2	4.7	6.6	7.0	6.7	4.5	5.7	6.0	7.4	4.9	6.1	5.2
الأطار B	4.1	4.9	6.2	6.9	6.8	4.4	5.7	5.8	6.9	4.9	6.0	4.9

هل يمكن القول أن الأطارات التي من النوع A تعطي أفضل أقتصاداً في البنزين من النوع B عندما $\alpha = 0.05$ ؟ استخدم اختبار الاشارة .

10 - في أحد المصانع التي يشتغل العمال بها لإنتاج سلعة معينة يدوياً أن هناك عمال ينتجون أكثر مما ينتجه الآخرين خلال اليوم الواحد ، ولكن هناك أعتقاد بأن بعض العمال الذين ينتجون أكثر أن صناعتهم لم تكن على مستوى عالي من الجودة ولتأكيد هذا الادعاء اختيرت عينة عشوائية من متوسط إنتاج 15 عاملاً خلال مدة شهر وتم حساب متوسط تقدير جودة الإنتاج لنفس العاملين فكانت النتائج كما يلي :

العامل	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
متوسط السلع المنتجة	12	15	35	21	20	17	19	46	20	25	39	25	30	27	29
متوسط تقدير الجودة	7.7	8.1	6.9	8.2	8.6	8.3	9.4	7.8	8.3	5.2	6.4	7.9	8.0	6.1	8.6

من هذه البيانات هل يمكن القول بوجود علاقة عكسية ما بين متوسط عدد القطع المنتجة ومتوسط تقدير جودتها عندما $\alpha = 0.05$ ؟

11 - من واقع زيارة مريض لطبيبة كان زمن أنتظاره قبل مقابلة الطبيب كما يلي :

17, 32, 25, 15, 15 , 28 , 25 , 20 , 12 , 35 , 20 , 26 , 24

من هذه البيانات هل يمكن القول أن المريض لاينتظر في المتوسط أكثر من 20 دقيقة حتى يقابل طبيبة عندما $\alpha = 0.05$ ؟ استخدم اختبار الاشارة .

12 - سئل 15 شخص مدمنين على الهيروين عن عمرهم الذى بداؤ به تعاطى المخدرات فكانت الأجابة كما يلي :

الشخص	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
العمر	22	24	37	28	15	14	22	16	18	17	23	16	20	18	15

من هذه البيانات هل يمكن القول بأن وسيط مجتمع المعاينة يختلف عن 20 عندما $\alpha = 0.10$ ؟ استخدم اختبار الاشارة .

13 - اجريت دراسة على معدل نبضات القلب عند الفئران عندما تكون وحيدة ومع بعضها البعض فكانت النتائج كما يلي :

الفئران	1	2	3	4	5	6	7	8	9	10
وحيد (x)	463	462	462	456	450	426	418	415	409	402
مع بعض (y)	523	494	461	535	476	454	448	408	470	437

من هذه البيانات هل يمكن القول أن وجود الفئران مع بعضها البعض يزيد من معدلات نبض القلب عندما $\alpha = 0.05$ ؟ استخدم اختبار الاشارة .

14 - بفرض أنه تم تصنيف 14 فني في مجال صيانة الالات بطريقة عشوائية للقيام بمهمتين ، حيث صنف 7 لكل مهمة وتم قياس الزمن الذي أستغرقه كل منهم لإنجاز المهمة فكانت النتائج كما يلي :

المهمة A	المهمة B
1.96	2.11
2.24	2.43
1.71	2.07
2.41	2.71
1.62	2.50
1.93	2.84
2.01	2.85

من هذه البيانات هل يمكن القول بأن هناك اختلاف ما بين التوزيع الاحتمالي لمجتمعي العينة عندما $\alpha = 0.05$ ؟ استخدم اختبار مان-وايتني .

15 - بفرض أن البيانات التالية تمثل التصنيف الناتج من إنتاج أحد خطوط الإنتاج لسلعة أستهلاكية بأحد المصانع :

D N N N N N N D D N N N N N N D D D N N N N N D N N N D D
N N N D D

حيث D تعني أن الانتاج غير قابل للاستهلاك و N تعني قابل للاستهلاك . من هذه البيانات هل يمكن القول بأن الانتاج عشوائي عندما $\alpha = 0.05$ ؟

16 - بطريقة عشوائية تم اختيار 12 عينة من إنتاج ورديتين للبلاستيك وثم قياس أقصى قوة (في كل 100 سم^2) بكل عينة فكانت النتائج كما يلي :

البلاستيك A : 15.8 , 16.3 , 18.9 , 17.6 , 21.4 , 16.9

البلاستيك B : 18.3 , 22.5 , 19.6 , 21.3 , 20.9 , 19.8

من هذه البيانات هل يمكن القول بوجود اختلاف ما بين قوة نوعي البلاستيك عندما $\alpha = 0.05$ ؟ استخدم اختبار مان-وايتني .

17 - صناعة معينة تتطلب مستوى عالي من المهارة ، ويعتقد أن الإنتاجية تزداد بزيادة سنوات الخبرة ، ولتأكيد هذا الاعتقاد اختيرت عينة عشوائية من 10 عاملين ممن يقومون بهذه الصناعة فكانت النتائج كالآتي :

العامل	1	2	3	4	5	6	7	8	9	10
الإنتاجية	4	6	10	2	12	6	5	10	13	9
سنوات الخبرة	80	82	88	81	92	85	83	86	91	90

هل هذه البيانات تشير إلى وجود ارتباط موجب ما بين سنوات الخبرة والإنتاجية ؟

18 - تدعى الشركة المنتجة لماكينات الحلاقة أن الماكينة الزوجية يمكن استخدامها العديد من المرات مقارنة بالماكينة المفردة ، ولتأكيد الادعاء تم اختيار عينتين عشوائيتين مستقلتين تحتوى كل منها على 8 ماكينات من كل نوع وسجل لكل منها عدد المرات التي استخدمت في الحلاقة قبل التخلص منها فكانت النتائج كما يلي :

الماكينة المفردة	الماكينة الزوجية
10	8
6	17
3	9
7	11
13	15
14	10
5	6
7	12

هل هذه البيانات تؤيد ادعاء الشركة عند $\alpha = 0.05$ ؟ استخدم اختبار مان-وايتني وما هو أفضل اختبار يمكن استخدامه لتحليل هذه البيانات ؟

19 - قام مجموعة من الأساتذة بدراسة على تلاميذ الصف الثالث الابتدائي من حيث هل هناك تغيير في رغبة التلاميذ في القراءة ، وذلك من خلال إعطائهم قصص وقراءتها بأنفسهم يومياً أو قراءة القصص عليهم ، وبطريقة عشوائية تم اختيار 20 تلميذ وتم تقسيمهم عشوائياً على البرنامجين وبعد فترة زمنية معينة طلب من مشرف كل برنامج إعطاء رتب من 1 (ضعيف) إلى 20 (ممتاز) لرغبة كل تلميذ في القراءة فكانت النتائج كما يلي :

القراءة عليهم	1	4	7	8	8	10	12	14	16	18
القراءة بأنفسهم	2	3	3	4	6	8	9	13	17	19

من هذه البيانات هل يمكن القول بوجود فروق معنوية ما بين المجموعتين عندما $\alpha = 0.05$ ؟ استخدم اختبار مان-وايتني .

20 - هناك اتجاه كبير في السنوات الأخيرة من قبل مصانع السيارات لإنتاج السيارات الصغيرة، ولمقارنة الشهرة بين أربعة أنواع من السيارات الصغيرة قام أحد الباحثين بجمع بيانات عن مبيعات السيارات من الأنواع الأربعة لمدة خمسة أشهر من أربعة مواقع للتسويق حيث كل موقع متخصص في بيع نوع معين فقط فكانت النتائج كما يلي :

	النـــــوع			
الشهر	A	B	C	D
1	9	17	14	8
2	10	20	16	9
3	13	15	19	12
4	11	12	19	11
5	7	18	13	8

من هذه البيانات هل هناك فروق بين عدد السيارات المباعة من كل نوع عندما $\alpha = 0.10$.

21 - تلعب درجة النعومة لبعض أنواع الورق دور في إقبال المستهلكين على شراؤها ، وإن إحدى الطرق المتبعة في تحديد درجة نعومة الورق هو أن يكون لذا المصنع أشخاص يحكمون على ذلك ، وبفرض أنه أعطى لعشرة أشخاص عينتين من إنتاجين وطلب منهم الحكم على درجة النعومة ، وسوف يكون الحكم من خلال أعطى رتب من 1 إلى 10 لكل إنتاج حيث أعلى ترتيب يعنى أن الإنتاج أكثر نعومة فكانت النتائج كما يلي :

الشخص	1	2	3	4	5	6	7	8	9	10
الإنتاج A	6	8	4	9	4	7	6	5	6	8
الإنتاج B	4	5	5	8	1	9	2	3	7	2

من هذه البيانات هل هناك اختلاف من درجتي النعومة للإنتاجين عندما $\alpha = 0.05$ ؟ استخدم اختبار ولكاكسن .

22 - بائع منسوجات يرغب في معرفة في ما إذا كان الأشخاص الذين يحملون مؤهلات علمية عالية لهم القدرة على البيع أكثر من الأشخاص الذين لا مؤهلات لهم . حيث تم اختيار عينة عشوائية من 8 أشخاص يحملون مؤهلات علمية و 12 شخص بدون مؤهلات وسجل مبيعات كل مجموعة خلال فترة زمنية معينة (المبيعات بالآلاف) فكانت النتائج كما يلي :

المبيعات (مؤهل)	67	103	96	97	76	84	76	102				
المبيعات(بدون مؤهل)	74	80	76	85	83	78	99	86	83	90	78	72

من هذه البيانات هل يمكن القول بأن وسيط مبيعات الأشخاص الذين يحملون مؤهلات علمية أكبر من وسيط مبيعات الأشخاص الذين لايحملون مؤهل علمي عندما $\alpha = 0.05$ ؟ استخدم اختبار مان-وايتني .

23 - أجريت تجربة لمقارنة الزمن المطلوب للشفاء من ثلاثة أنواع من الأنفلونزا هي A ، B ، C حيث تم اختيار عينة عشوائية تتكون من 21 من مجموعة من المتبرعين وتم تقسيمهم إلى ثلاث مجموعات كل مجموعة بها 7 أشخاص وكل مجموعة تم تصنيفها بطريقة عشوائية لنوع معين من الأنفلونزا وتم حقن كل مجموعة بالأنفلونزا المصنفة لها ثم تمت معالجة جميع الأشخاص تحت نفس الظروف وسجل الزمن الذي أستغرقه كل شخص للشفاء من المرض (الزمن بالأيام) فكانت النتائج كما يلي :

نوع الأنفلونزا

A	B	C
12	9	7
6	10	3
13	5	7
10	4	5
8	9	6
11	8	4
7	11	8

هل هذه البيانات تشير إلى أن الزمن المطلوب للشفاء من نوع أو اكثر من الأنواع أكبر من الأنواع الأخرى عندما $\alpha = 0.05$ ؟ استخدم اختبار كروسكل .

24 - شركة التسويق تقوم بالدعاية لمبيعاتها عن طريق المراسلة والجرائد والمجلات من خلال 15 قناة توزيع منتشرة على كامل البلاد وقد سجلت نسبة الزبائن الذين اشتروا منها عن طريق الدعاية خلال فترة زمنية مدتها سنة فكانت النتائج كما يلي :

الموقع (القناة)	المراسلة	الجرائد	المجلات
1	7.3	15.7	10.1
2	9.4	18.3	8.2
3	4.3	11.2	5.1
4	11.3	19.1	6.5
5	3.3	9.2	8.7
6	4.2	10.5	6.0
7	5.9	8.7	12.3
8	6.2	14.3	11.1
9	4.3	3.1	6.0
10	10.0	18.8	12.1
11	2.2	5.7	6.3
12	6.3	20.2	4.3
13	8.0	14.1	9.1
14	7.4	6.2	18.1
15	3.2	8.9	5.0

من هذه البيانات هل يمكن القول بأن هناك فرق في الاستجابة بين الأنواع الثلاثة من الدعاية عندما $\alpha = 0.01$ ؟ (نسبة الاستجابة = عدد الزبائن الذين اشتروا بناءً على نوع الدعاية على إجمالي الزبائن الذين اشتروا من ذلك الموقع) استخدم اختبار فريدمان ؟

25 - أجريت تجربة لمقارنة تأثير السم الناتج من ثلاثة أنواع من المركبات الكيميائية على جلد الفئران ، حيث وضعت ثلاثة علامات مربعة بجوار بعضها البعض على ظهر كل فأر ، وتم وضع الأنواع الكيميائية الثلاثة على كل فأر ثم أعطيت درجة من 0 إلى 10 على كل مربع على ظهر كل فأر وذلك اعتماداً على شدة الإصابة فكانت النتائج كما يلي :

المـــركب الكيميائي

الفأر	A	B	C
1	6	5	2
2	9	8	4
3	6	9	3
4	5	8	6
5	7	.8	9
6	5	7	6
7	6	7	5
8	6	5	7

من هذه البيانات هل يمكن القول بأن تأثير الأنواع الكيميائية الثلاثة مختلف عندما $\alpha = 0.01$ ؟ استخدم اختبار فريدمان .

26 - في تجربة للتحكم في بيئة معمل بقسم علم الحيوان تم اختيار عشرة رجال وعشرة نساء لتحديد درجات الحرارة التي يرون أنها مناسبة فكانت النتائج كما يلي :

الرجال	74	72	77	76	76	73	75	73	74	75
النساء	75	77	78	79	77	73	78	79	78	80

على افتراض أن هذه البيانات تمثل عينتين عشوائيتين من مجتمعيهما، هل يمكن القول بأن متوسطي درجة الحرارة المناسبة للرجال والنساء متساوية عندما $\alpha = 0.05$ ؟ استخدم اختبار مان-وايتني .

27 - على مدير إدارة أحد المصانع الاختيار ما بين طريقتين هما A و B لتحسين احتياطات الأمان للعاملين به ، وللوصول لقرار للاختيار بينهما تم اختبار الطريقتين من قبل عاملين متخصصين في مجال السلامة وطلب من كل منهم تقييم الطريقتين بمقياس من 1 إلى 10 (حيث التقدير الأعلى يعني الطريقة أفضل) ، وسوف يعتمد مدير الإدارة الطريقة B إذا كانت تقديرات رجال السلامة لها أعلى من تقديراتهم للطريقة A وذلك لأن هذه الأخيرة تكاليف تنفيذها عالية ، فكانت نتائج الدراسة كما يلي :

الخبراء	1	2	3	4	5	6	7	8	9	10
الطريقة A	7	4	8	9	3	6	8	10	9	9
الطريقة B	9	5	8	8	6	10	9	8	4	5

من هذه البيانات هل يمكن القول بأن الطريقة B أفضل من الطريقة A عندما $\alpha = 0.05$ ؟ استخدم اختبار ولكاكسن .

28 - لاكتشاف إمكانية وجود فروق في معدلات الإنتاج بين ثلاثة خطوط تنتج نفس النوع من السلعة ، تم اختيار عينات عشوائية مستقلة من إجمال الإنتاج ولمدة سبعة أيام من كل خط فكانت البيانات كما يلي :

الخط		
1	2	3
48	41	18
43	36	42
39	29	28
57	40	38
21	35	15
47	45	33
58	32	31

هل هذه البيانات تشير إلى فروق في أجمالي الإنتاج بين خطوط الإنتاج عند مستوى المعنوية 5٪ ؟ استخدم اختبار كروسكل .

29 - أجريت تجربة لمقارنة ثلاثة أنواع من الآلات الحاسبة A ، B ، C من حيث أيها أسهل استعمالاً ، ولعمل هذه المقارنة تم اختيار ستة طلاب بطريقة عشوائية وطلب منهم القيام بمتابعة من العمليات الحسابية على كل آلة من الآلات الثلاث . إن استعمال الآلات من قبل الطلاب تم بطريقة عشوائية وسجل الزمن المطلوب لإنهاء العمليات الحسابية من قبل كل طالب (بالثواني) وذلك كما يلي :

الطالب	A	B	C
1	306	330	300
2	260	265	285
3	281	290	277
4	288	301	305
5	301	309	319
6	262	245	240

من هذه البيانات هل يمكن القول بأن هناك اختلاف بين الآلات الثلاثة من حيث سهولة الاستعمال عندما $\alpha = 0.05$ ؟ استخدم اختبار فريدمان .

30 - تقدم عشرة أشخاص لشغل وظيفة معينة وتم ترتيبهم من 1 (الأفضل) إلى 10 (الأقل) من قبل شخصين متخصصين هما A و B في نفس المجال فكانت النتائج كما يلي :

	المتقدم									
المتخصص	1	2	3	4	5	6	7	8	9	10
A	4	1	9	5	2	10	7	3	6	8
B	5	2	10	6	1	9	7	3	4	8

من هذه البيانات هل يمكن القول بوجود علاقة ما بين الرتب المعطاة من قبل الشخصين ؟

31 - ترغب إدارة أحد المصانع في تحديد ما إذا كان عدد القطع المنتجة وبها عيوب يزداد كلما تقدم اليوم ، وبدون معرفة العاملين قامت الإدارة بالكشف عن كل قطعة يتم إنتاجها يومياً وسجلت نسبة القطع المعيبة فكانت النتائج كما يلي :

الساعة	1	2	3	4	5	6	7	8
نسبة العيب	0.10	0.11	0.09	0.06	0.08	0.03	0.05	0.02

من هذه البيانات هل يمكن القول بأن نسبة القطع المنتجة وبها عيوب تزداد كلما تقدم اليوم عندما $\alpha = 0.05$ ؟

32 - يدعى مدير إدارة أحدى الشركات انه يتم تعيين الموظفين بطريقة عشوائية وذلك بصرف النظر عن الجنس ، ولتأكيد أدعاءه أختيرت عينة عشوائية من 13 موظفاً تم تعيينهم مؤخراً فكان كما يلي :

M M M M F M M M F F M F F

حيث F تعنى أنثى و تعني M ذكر . من هذه البيانات هل يمكن القول بأن أدعاء المدير صحيحاً عندما $\alpha = 0.05$ ؟

ملحق الجداول الإحصائية

1 - جدول (1) : الارقام العشوائية .
2 - جدول (2) : التوزيع الاحتمالي لذي الحدين .
3 - جدول (3) : التوزيع الاحتمالي لبواسون .
4 - جدول (4) : التوزيع الطبيعي المعياري .
5 - جدول (5) : القيم المئوية لتوزيع t .
6 - جدول (6) : القيم المئوية لتوزيع مربع كاي .
7 - جدول (7) : القيم المئوية لتوزيع F .
8 - جدول (8) : مئويات اختبار مان – وايتني .
9 - جدول (9) : مئويات اختبار رتب الإشارة ولكاكسن .
10 - جدول (10) : مئويات اختبار كروسكل – وليس .
11 - جدول (11) : مئويات اختبار كولو مجروف – سميرنوف .
12 - جدول (12) : مئويات اختبار سبيرمان .
13 - جدول (13) : التحويل من r إلى z .
14 - جدول (14) : القيم الجدولية لاختبار العشوائية .
15 - جدول (15) : القيم الجدولية لاختبار دونيت .
16 - جدول (16) : القيم المئوية لاختبار توكي .

جدول (1) : الارقام العشوائية

2671	4690	1550	2262	2597	8034	0785	2978	4409	0237
9111	0250	3275	7519	9740	4577	2064	0286	3398	1348
0391	6035	9230	4999	3332	0608	6113	0391	5789	9926
2475	2144	1886	2079	3004	9686	5669	4367	9306	2595
5336	5845	2095	6446	5694	3641	1085	8705	5416	9066
6808	0423	0155	1652	7897	4335	3567	7109	9690	3739
8525	0577	8940	9451	6726	0876	3818	7607	8854	3566
0398	0741	8787	3043	5063	0617	1770	5048	7721	7032
3623	9636	3638	1406	5731	3978	8068	7238	9715	3363
0739	2644	4917	8866	3632	5399	5175	7422	2476	2607
6713	3041	8133	8749	8835	6745	3597	3476	3816	3455
7775	9315	0432	8327	0861	1515	2297	3375	3713	9174
8599	2122	6842	9202	0810	2936	1514	2090	3067	3574
7955	3759	5254	1126	5553	4713	9605	7909	1658	5490
4766	0070	7260	6033	7997	0109	5993	7592	5436	1727
5165	1670	2534	8811	8231	3721	7947	5719	2640	1394
9111	0513	2751	8256	2931	7783	1281	6531	7259	6993
1667	1084	7889	8963	7018	8617	6381	0723	4926	4551
2145	4587	8585	2412	5431	4667	1942	7238	9613	2212
2739	5528	1481	7528	9368	1823	6979	2547	7268	2467
8769	5480	9160	5354	9700	1362	2774	7980	9157	8788
6531	9435	3422	2474	1475	0159	3414	5224	8399	5820
2937	4134	7120	2206	5084	9473	3958	7320	9878	8609
1581	3285	3727	8924	6204	0797	0882	5945	9375	9153
6268	1045	7076	1436	4165	0143	0293	4190	7171	7932
4293	0523	8625	1961	1039	2856	4889	4358	1492	3804
6936	4213	3212	7229	1230	0019	5998	9206	6753	3762
5334	7641	3258	3769	1362	2771	6124	9813	7915	8960
9373	1158	4418	8826	5665	5896	0358	4717	8232	4859
6968	9428	8950	5346	1741	2348	8143	5377	7695	0685
4229	0587	8794	4009	9691	4579	3302	7673	9629	5246
3807	7785	7097	5701	6639	0723	4819	0900	2711	7650
4891	8829	1642	2155	0796	0466	2946	2970	9143	6590
1055	2968	7911	7479	8199	9735	8271	5339	7058	2964
2983	2345	0568	4125	0894	8302	0506	6761	7706	4310
4026	3129	2968	8053	2797	4022	9838	9611	9975	2437
4075	0260	4256	0337	2355	9371	2954	6021	5783	2827
8488	5450	1327	7358	2034	8060	1788	6913	6123	9405
1976	1749	5742	4098	5887	4567	6064	2777	7830	5668
2793	4701	9466	9554	8294	2160	7486	1557	4769	2781

المصدر : Handbook of Statistical Tables Addison- wesley co., 1962

تابع جدول (1)

0916	6272	6825	7188	9611	1181	2301	5516	5451	6832
5961	1149	7946	1950	2010	0600	5655	0796	0569	4365
3222	4189	1891	8172	8731	4769	2782	1325	4238	9279
1176	7834	4600	9992	9449	5824	5344	1008	6678	1921
2369	8971	2314	4806	5071	8908	8274	4936	3357	4441
0041	4329	9265	0352	4764	9070	7527	7791	1094	2008
0803	8302	6814	2422	6351	0637	0514	0246	1845	8594
9965	7804	3930	8803	0268	1426	3130	3613	3947	8086
0011	2387	3148	7559	4216	2946	2865	6333	1916	2259
1767	9871	3914	5790	5287	7915	8959	1346	5482	9251
2604	3074	0504	3828	7881	0797	1094	4098	4940	7067
6930	4180	3074	0060	0909	3187	8991	0682	2385	2307
6160	9899	9084	5704	5666	3051	0325	4733	5905	9226
4884	1857	2847	2581	4870	1782	2980	0587	8797	5545
7294	2009	9020	0006	4309	3941	5645	6238	5052	4150
3478	4973	1056	3687	3145	5988	4214	5543	9185	9375
1764	7860	4150	2881	9895	2531	7363	8756	3724	9359
3025	0890	6436	3461	1411	0303	7422	2684	6256	3495
1771	3056	6630	4982	2386	2517	4747	5505	8785	8708
0254	1892	9066	4890	8716	2258	2452	3913	6790	6331
8537	9966	8224	9151	1855	8911	4422	1913	2000	1482
1475	0261	4465	4803	8231	6469	9935	4256	0648	7768
5209	5769	8410	3041	4325	7290	3381	5209	5571	9458
5456	5944	6038	3210	7165	0723	4820	1846	0005	3865
8043	6694	4853	8425	5871	1322	1052	1452	2486	1669
1719	0148	6977	1244	6443	5955	7945	1218	9391	6485
7432	2955	3933	8110	8585	1893	9218	7153	7566	6040
4926	4761	7812	7439	6436	3145	5934	7852	9095	9497
0769	0683	3768	1048	8519	2987	0124	3064	1881	3177
0865	3139	8514	5014	3274	6395	0549	3858	0820	6406
0204	7273	4964	5475	2648	6977	1371	6971	4850	6873
0692	1733	2349	2648	6609	5676	6445	3271	8867	3469
3139	4861	3066	9783	5068	4852	4143	7923	3858	0504
2033	7430	4389	7121	9982	0651	9119	9731	6421	4731
3921	0530	3665	8455	4205	7363	3084	3931	9311	1313
4711	9244	8135	9877	9529	9160	4407	9077	5306	0054
[illegible]	1570	6654	3616	2949	7001	5185	7108	9270	6550
6715	8079	6584	4319	9361	1087	9549	1634	4808	3686
3790	2196	3743	1909	3129	1828	8994	6034	5656	3035
7212	9909	[illegible]	7660	2670	6406	0690	4240	4070	6549

تابع جدول (1)

6701	0154	8806	1716	7029	6776	9465	8818	2886	3547
3777	9532	1333	8131	2929	6987	2408	0487	9172	6177
2495	3054	1692	0089	4090	2983	2136	8947	4625	7177
2073	8878	9742	3012	0042	3996	9930	1651	4982	9645
2252	8004	7840	2105	3033	8749	9153	2872	5100	8674
2104	2224	4052	2273	4753	4505	7156	5417	9725	7599
2371	0005	3844	6654	3246	4853	4301	8886	5217	1153
3270	1214	9649	1872	6930	9791	0248	2687	8126	1501
6209	7237	1966	5541	4224	7080	7630	6422	1160	5675
1309	9126	2920	4359	1726	0562	9654	4182	4097	7493
2406	8013	3634	6428	8091	5925	3923	1686	6097	9670
7365	9859	9378	7084	9402	9201	1815	7064	4324	7081
2889	4738	9929	1476	0785	3832	1281	5821	3690	9185
7951	3781	4755	6986	1659	5727	8108	9816	5759	4188
4548	6778	7672	9101	3911	8127	1918	8512	4197	6402
5701	8342	2852	4278	3343	9830	1756	0546	6717	3114
2187	7266	1210	3797	1636	7917	9933	3518	6923	6349
9360	6640	1315	6284	8265	7232	0291	3467	1088	7834
7850	7626	0745	1992	4998	7349	6451	6186	8916	4292
6186	9233	6571	0925	1748	5490	5264	3820	9829	1335

جدول (2) : التوزيع الاحتمالي لذي الحدين لقيم مختلفة لكل من (n , p) ، حيث

$$p_X(r) = P(X = r) = \binom{n}{r} p^r q^{n-r} \qquad r = 0, 1, 2, \cdots, n$$

n = 1

r \ p	.01	.02	.03	.04	.05	.06	.07	.08	.09	.10
0	.9900	.9800	.9700	.9600	.9500	.9400	.9300	.9200	.9100	.9000
1	.0100	.0200	.0300	.0400	.0500	.0600	.0700	.0800	.0900	.1000
	.11	.12	.13	.14	.15	.16	.17	.18	.19	.20
0	.8900	.8800	.8700	.8600	.8500	.8400	.8300	.8200	.8100	.8000
1	.1100	.1200	.1300	.1400	.1500	.1600	.1700	.1800	.1900	.2000
	.21	.22	.23	.24	.25	.26	.27	.28	.29	.30
0	.7900	.7800	.7700	.7600	.7500	.7400	.7300	.7200	.7100	.7000
1	.2100	.2200	.2300	.2400	.2500	.2600	.2700	.2800	.2900	.3000
	.31	.32	.33	.34	.35	.36	.37	.38	.39	.40
0	.6900	.6800	.6700	.6600	.6500	.6400	.6300	.6200	.6100	.6000
1	.3100	.3200	.3300	.3400	.3500	.3600	.3700	.3800	.3900	.4000
	.41	.42	.43	.44	.45	.46	.47	.48	.49	.50
0	.5900	.5800	.5700	.5600	.5500	.5400	.5300	.5200	.5100	.5000
1	.4100	.4200	.4300	.4400	.4500	.4600	.4700	.4800	.4900	.5000

n = 2

r \ p	.01	.02	.03	.04	.05	.06	.07	.08	.09	.10
0	.9801	.9604	.9409	.9216	.9025	.8836	.8649	.8464	.8281	.8100
1	.0198	.0392	.0582	.0768	.0950	.1128	.1302	.1472	.1638	.1800
2	.0001	.0004	.0009	.0016	.0025	.0036	.0049	.0064	.0081	.0100
	.11	.12	.13	.14	.15	.16	.17	.18	.19	.20
0	.7921	.7744	.7569	.7396	.7225	.7056	.6889	.6724	.6561	.6400
1	.1958	.2112	.2262	.2408	.2550	.2688	.2822	.2952	.3078	.3200
2	.0121	.0144	.0169	.0196	.0225	.0256	.0289	.0324	.0361	.0400
	.21	.22	.23	.24	.25	.26	.27	.28	.29	.30
0	.6241	.6084	.5929	.5776	.5625	.5476	.5329	.5184	.5041	.4900
1	.3318	.3432	.3542	.3648	.3750	.3848	.3942	.4032	.4118	.4200
2	.0441	.0484	.0529	.0576	.0625	.0676	.0729	.0784	.0841	.0900
	.31	.32	.33	.34	.35	.36	.37	.38	.39	.40
0	.4761	.4624	.4489	.4356	.4225	.4096	.3969	.3844	.3721	.3600
1	.4278	.4352	.4422	.4488	.4550	.4608	.4662	.4712	.4758	.4800
2	.0961	.1024	.1089	.1156	.1225	.1296	.1369	.1444	.1521	.1600
	.41	.42	.43	.44	.45	.46	.47	.48	.49	.50
0	.3481	.3364	.3249	.3136	.3025	.2916	.2809	.2704	.2601	.2500
1	.4838	.4872	.4902	.4928	.4950	.4968	.4982	.4992	.4998	.5000
2	.1681	.1764	.1849	.1936	.2025	.2116	.2209	.2304	.2401	.2500

Wayne W. Daniel , Applied Nonparametric Statistics , (1978) , H.M. co. : المصدر

جدول (2)

n = 3

x \ p	.01	.02	.03	.04	.05	.06	.07	.08	.09	.10
0	9703	9412	9127	8847	8574	8306	8044	7787	7536	7290
1	0294	0576	0847	1106	1354	1590	1816	2031	2236	2430
2	0003	0012	0026	0046	0071	0102	0137	0177	0221	0270
3	0000	0000	0000	0001	0001	0002	0003	0005	0007	0010

x \ p	.11	.12	.13	.14	.15	.16	.17	.18	.19	.20
0	7050	6815	6585	6361	6141	5927	5718	5514	5314	5120
1	2614	2788	2952	3106	3251	3387	3513	3631	3740	3840
2	0323	0380	0441	0506	0574	0645	0720	0797	0877	0960
3	0013	0017	0022	0027	0034	0041	0049	0058	0069	0080

x \ p	.21	.22	.23	.24	.25	.26	.27	.28	.29	.30
0	4930	4746	4565	4390	4219	4052	3890	3732	3579	3430
1	3932	4015	4091	4159	4219	4271	4316	4355	4386	4410
2	1045	1133	1222	1313	1406	1501	1597	1693	1791	1890
3	0093	0106	0122	0138	0156	0176	0197	0220	0244	0270

x \ p	.31	.32	.33	.34	.35	.36	.37	.38	.39	.40
0	3285	3144	3008	2875	2746	2621	2500	2383	2270	2160
1	4428	4439	4444	4443	4436	4424	4406	4382	4354	4320
2	1989	2089	2189	2289	2389	2488	2587	2686	2783	2880
3	0298	0328	0359	0393	0429	0467	0507	0549	0593	0640

x \ p	.41	.42	.43	.44	.45	.46	.47	.48	.49	.50
0	2054	1951	1852	1756	1664	1575	1489	1406	1327	1250
1	4282	4239	4191	4140	4084	4024	3961	3894	3823	3750
2	2975	3069	3162	3252	3341	3428	3512	3594	3674	3750
3	0689	0741	0795	0852	0911	0973	1038	1106	1176	1250

n = 4

x \ p	.01	.02	.03	.04	.05	.06	.07	.08	.09	.10
0	9606	9224	8853	8493	8145	7807	7481	7164	6857	6561
1	0388	0753	1095	1416	1715	1993	2252	2492	2713	2916
2	0006	0023	0051	0088	0135	0191	0254	0325	0402	0486
3	0000	0000	0001	0002	0005	0008	0013	0019	0027	0036
4	0000	0000	0000	0000	0000	0000	0000	0000	0001	0001

x \ p	.11	.12	.13	.14	.15	.16	.17	.18	.19	.20
0	6274	5997	5728	5470	5220	4979	4746	4521	4305	4096
1	3102	3271	3424	3562	3685	3793	3888	3970	4039	4096
2	0575	0669	0767	0870	0975	1084	1195	1307	1421	1536
3	0047	0061	0076	0094	0115	0138	0163	0191	0222	0256
4	0001	0002	0003	0004	0005	0007	0008	0010	0013	0016

x \ p	.21	.22	.23	.24	.25	.26	.27	.28	.29	.30
0	3895	3702	3515	3336	3164	2999	2840	2687	2541	2401
1	4142	4176	4200	4214	4219	4214	4201	4180	4152	4116
2	1651	1767	1882	1996	2109	2221	2331	2439	2544	2646
3	0293	0332	0375	0420	0469	0520	0575	0632	0693	0756
4	0019	0023	0028	0033	0039	0046	0053	0061	0071	0081

x \ p	.31	.32	.33	.34	.35	.36	.37	.38	.39	.40
0	2267	2138	2015	1897	1785	1678	1575	1478	1385	1296
1	4074	4025	3970	3910	3845	3775	3701	3623	3541	3456
2	2745	2841	2933	3021	3105	3185	3260	3330	3396	3456
3	0822	0891	0963	1038	1115	1194	1276	1361	1447	1536
4	0092	0105	0119	0134	0150	0168	0187	0208	0231	0256

x \ p	.41	.42	.43	.44	.45	.46	.47	.48	.49	.50
0	1212	1132	1056	0983	0915	0850	0789	0731	0677	0625
1	3368	3278	3185	3091	2995	2897	2799	2700	2600	2500
2	3511	3560	3604	3643	3675	3702	3723	3738	3747	3750
3	1627	1719	1813	1908	2005	2102	2201	2300	2400	2500
4	0283	0311	0342	0375	0410	0448	0488	0531	0576	0625

تابع جدول (2)

$n = 5$

r \ p	.01	.02	.03	.04	.05	.06	.07	.08	.09	.10
0	.9510	.9039	.8587	.8154	.7738	.7339	.6957	.6591	.6240	.5905
1	.0480	.0922	.1328	.1699	.2036	.2342	.2618	.2866	.3086	.3280
2	.0010	.0038	.0082	.0142	.0214	.0299	.0394	.0498	.0610	.0729
3	.0000	.0001	.0003	.0006	.0011	.0019	.0030	.0043	.0060	.0081
4	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0002	.0003	.0004

r \ p	.11	.12	.13	.14	.15	.16	.17	.18	.19	.20
0	.5584	.5277	.4984	.4704	.4437	.4182	.3939	.3707	.3487	.3277
1	.3451	.3598	.3724	.3829	.3915	.3983	.4034	.4069	.4089	.4096
2	.0853	.0981	.1113	.1247	.1382	.1517	.1652	.1786	.1919	.2048
3	.0105	.0134	.0166	.0203	.0244	.0289	.0338	.0392	.0450	.0512
4	.0007	.0009	.0012	.0017	.0022	.0028	.0035	.0043	.0053	.0064
5	.0000	.0000	.0000	.0001	.0001	.0001	.0001	.0002	.0002	.0003

r \ p	.21	.22	.23	.24	.25	.26	.27	.28	.29	.30
0	.3077	.2887	.2707	.2536	.2373	.2219	.2073	.1935	.1804	.1681
1	.4090	.4072	.4043	.4003	.3955	.3898	.3834	.3762	.3685	.3602
2	.2174	.2297	.2415	.2529	.2637	.2739	.2836	.2926	.3010	.3087
3	.0578	.0648	.0721	.0798	.0879	.0962	.1049	.1138	.1229	.1323
4	.0077	.0091	.0108	.0126	.0146	.0169	.0194	.0221	.0251	.0284
5	.0004	.0005	.0006	.0008	.0010	.0012	.0014	.0017	.0021	.0024

r \ p	.31	.32	.33	.34	.35	.36	.37	.38	.39	.40
0	.1564	.1454	.1350	.1252	.1160	.1074	.0992	.0916	.0845	.0778
1	.3513	.3421	.3325	.3226	.3124	.3020	.2914	.2808	.2700	.2592
2	.3157	.3220	.3275	.3323	.3364	.3397	.3423	.3441	.3452	.3456
3	.1418	.1515	.1613	.1712	.1811	.1911	.2010	.2109	.2207	.2304
4	.0319	.0357	.0397	.0441	.0488	.0537	.0590	.0646	.0706	.0768
5	.0029	.0034	.0039	.0045	.0053	.0060	.0069	.0079	.0090	.0102

r \ p	.41	.42	.43	.44	.45	.46	.47	.48	.49	.50
0	.0715	.0656	.0602	.0551	.0503	.0459	.0418	.0380	.0345	.0312
1	.2484	.2376	.2270	.2164	.2059	.1956	.1854	.1755	.1657	.1562
2	.3452	.3442	.3424	.3400	.3369	.3332	.3289	.3240	.3185	.3125
3	.2399	.2492	.2583	.2671	.2757	.2838	.2916	.2990	.3060	.3125
4	.0834	.0902	.0974	.1049	.1128	.1209	.1293	.1380	.1470	.1562
5	.0116	.0131	.0147	.0165	.0185	.0206	.0229	.0255	.0282	.0312

$n = 6$

r \ p	.01	.02	.03	.04	.05	.06	.07	.08	.09	.10
0	.9415	.8858	.8330	.7828	.7351	.6899	.6470	.6064	.5679	.5314
1	.0571	.1085	.1546	.1957	.2321	.2642	.2922	.3164	.3370	.3543
2	.0014	.0055	.0120	.0204	.0305	.0422	.0550	.0688	.0833	.0984
3	.0000	.0002	.0005	.0011	.0021	.0036	.0055	.0080	.0110	.0146
4	.0000	.0000	.0000	.0000	.0001	.0002	.0003	.0005	.0008	.0012
5	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001

r \ p	.11	.12	.13	.14	.15	.16	.17	.18	.19	.20
0	.4970	.4644	.4336	.4046	.3771	.3513	.3269	.3040	.2824	.2621
1	.3685	.3800	.3888	.3952	.3993	.4015	.4018	.4004	.3975	.3932
2	.1139	.1295	.1452	.1608	.1762	.1912	.2057	.2197	.2331	.2458
3	.0188	.0236	.0289	.0349	.0415	.0486	.0562	.0643	.0729	.0819
4	.0017	.0024	.0032	.0043	.0055	.0069	.0086	.0106	.0128	.0154
5	.0001	.0001	.0002	.0003	.0004	.0005	.0007	.0009	.0012	.0015
6	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001

r \ p	.21	.22	.23	.24	.25	.26	.27	.28	.29	.30
0	.2431	.2252	.2084	.1927	.1780	.1642	.1513	.1393	.1281	.1176
1	.3877	.3811	.3735	.3651	.3560	.3462	.3358	.3251	.3139	.3025
2	.2577	.2687	.2789	.2882	.2966	.3041	.3105	.3160	.3206	.3241
3	.0913	.1011	.1111	.1214	.1318	.1424	.1531	.1639	.1746	.1852
4	.0182	.0214	.0249	.0287	.0330	.0375	.0425	.0478	.0535	.0595
5	.0019	.0024	.0030	.0036	.0044	.0053	.0063	.0074	.0087	.0102
6	.0001	.0001	.0001	.0002	.0002	.0003	.0004	.0005	.0006	.0007

تابع جدول (2)

r \ p	.31	.32	.33	.34	.35	.36	.37	.38	.39	.40
					n = 6					
0	.1079	.0989	.0905	.0827	.0754	.0687	.0625	.0568	.0515	.0467
1	.2909	.2792	.2673	.2555	.2437	.2319	.2203	.2089	.1976	.1866
2	.3267	.3284	.3292	.3290	.3280	.3261	.3235	.3201	.3159	.3110
3	.1957	.2061	.2162	.2260	.2355	.2446	.2533	.2616	.2693	.2765
4	.0660	.0727	.0799	.0873	.0951	.1032	.1116	.1202	.1291	.1382
5	.0119	.0137	.0157	.0180	.0205	.0232	.0262	.0295	.0330	.0369
6	.0009	.0011	.0013	.0015	.0018	.0022	.0026	.0030	.0035	.0041

r \ p	.41	.42	.43	.44	.45	.46	.47	.48	.49	.50
0	.0422	.0381	.0343	.0308	.0277	.0248	.0222	.0198	.0176	.0156
1	.1759	.1654	.1552	.1454	.1359	.1267	.1179	.1095	.1014	.0938
2	.3055	.2994	.2928	.2856	.2780	.2699	.2615	.2527	.2436	.2344
3	.2831	.2891	.2945	.2992	.3032	.3065	.3091	.3110	.3121	.3125
4	.1475	.1570	.1666	.1763	.1861	.1958	.2056	.2153	.2249	.2344
5	.0410	.0455	.0503	.0554	.0609	.0667	.0729	.0795	.0864	.0938
6	.0048	.0055	.0063	.0073	.0083	.0095	.0108	.0122	.0138	.0156

r \ p	.01	.02	.03	.04	.05	.06	.07	.08	.09	.10
					n = 7					
0	.9321	.8681	.8080	.7514	.6983	.6485	.6017	.5578	.5168	.4783
1	.0659	.1240	.1749	.2192	.2573	.2897	.3170	.3396	.3578	.3720
2	.0020	.0076	.0162	.0274	.0406	.0555	.0716	.0886	.1061	.1240
3	.0000	.0003	.0008	.0019	.0036	.0059	.0090	.0128	.0175	.0230
4	.0000	.0000	.0000	.0001	.0002	.0004	.0007	.0011	.0017	.0026
5	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0002

r \ p	.11	.12	.13	.14	.15	.16	.17	.18	.19	.20
0	.4423	.4087	.3773	.3479	.3206	.2951	.2714	.2493	.2288	.2097
1	.3827	.3901	.3946	.3965	.3960	.3935	.3891	.3830	.3756	.3670
2	.1419	.1596	.1769	.1936	.2097	.2248	.2391	.2523	.2643	.2753
3	.0292	.0363	.0441	.0525	.0617	.0714	.0816	.0923	.1033	.1147
4	.0036	.0049	.0066	.0086	.0109	.0136	.0167	.0203	.0242	.0287
5	.0003	.0004	.0006	.0008	.0012	.0016	.0021	.0027	.0034	.0043
6	.0000	.0000	.0000	.0000	.0001	.0001	.0001	.0002	.0003	.0004

r \ p	.21	.22	.23	.24	.25	.26	.27	.28	.29	.30
0	.1920	.1757	.1605	.1465	.1335	.1215	.1105	.1003	.0910	.0824
1	.3573	.3468	.3356	.3237	.3115	.2989	.2860	.2731	.2600	.2471
2	.2850	.2935	.3007	.3067	.3115	.3150	.3174	.3186	.3186	.3177
3	.1263	.1379	.1497	.1614	.1730	.1845	.1956	.2065	.2169	.2269
4	.0336	.0389	.0447	.0510	.0577	.0648	.0724	.0803	.0886	.0972
5	.0054	.0066	.0080	.0097	.0115	.0137	.0161	.0187	.0217	.0250
6	.0005	.0006	.0008	.0010	.0013	.0016	.0020	.0024	.0030	.0036
7	.0000	.0000	.0000	.0000	.0001	.0001	.0001	.0001	.0002	.0002

r \ p	.31	.32	.33	.34	.35	.36	.37	.38	.39	.40
0	.0745	.0672	.0606	.0546	.0490	.0440	.0394	.0352	.0314	.0280
1	.2342	.2215	.2090	.1967	.1848	.1732	.1619	.1511	.1407	.1306
2	.3156	.3127	.3088	.3040	.2985	.2922	.2853	.2778	.2698	.2613
3	.2363	.2452	.2535	.2610	.2679	.2740	.2793	.2838	.2875	.2903
4	.1062	.1154	.1248	.1345	.1442	.1541	.1640	.1739	.1838	.1935
5	.0286	.0326	.0369	.0416	.0466	.0520	.0578	.0640	.0705	.0774
6	.0043	.0051	.0061	.0071	.0084	.0098	.0113	.0131	.0150	.0172
7	.0003	.0003	.0004	.0005	.0006	.0008	.0009	.0011	.0014	.0016

r \ p	.41	.42	.43	.44	.45	.46	.47	.48	.49	.50
0	.0249	.0221	.0195	.0173	.0152	.0134	.0117	.0103	.0090	.0078
1	.1211	.1119	.1032	.0950	.0872	.0798	.0729	.0664	.0604	.0547
2	.2524	.2431	.2336	.2239	.2140	.2040	.1940	.1840	.1740	.1641
3	.2923	.2934	.2937	.2932	.2918	.2897	.2867	.2830	.2786	.2734
4	.2031	.2125	.2216	.2304	.2388	.2468	.2543	.2612	.2676	.2734
5	.0847	.0923	.1003	.1086	.1172	.1261	.1353	.1447	.1543	.1641
6	.0196	.0223	.0252	.0284	.0320	.0358	.0400	.0445	.0494	.0547
7	.0019	.0023	.0027	.0032	.0037	.0044	.0051	.0059	.0068	.0078

تابع جدول (2)

n = 8

r \ p	.01	.02	.03	.04	.05	.06	.07	.08	.09	.10
0	.9227	.8508	.7837	.7214	.6634	.6096	.5596	.5132	.4703	.4305
1	.0746	.1389	.1939	.2405	.2793	.3113	.3370	.3570	.3721	.3826
2	.0026	.0099	.0210	.0351	.0515	.0695	.0888	.1087	.1288	.1488
3	.0001	.0004	.0013	.0029	.0054	.0089	.0134	.0189	.0255	.0331
4	.0000	.0000	.0001	.0002	.0004	.0007	.0013	.0021	.0031	.0046
5	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0002	.0004

r \ p	.11	.12	.13	.14	.15	.16	.17	.18	.19	.20
0	.3937	.3596	.3282	.2992	.2725	.2479	.2252	.2044	.1853	.1678
1	.3892	.3923	.3923	.3897	.3847	.3777	.3691	.3590	.3477	.3355
2	.1684	.1872	.2052	.2220	.2376	.2518	.2646	.2758	.2855	.2936
3	.0416	.0511	.0613	.0723	.0839	.0959	.1084	.1211	.1339	.1468
4	.0064	.0087	.0115	.0147	.0185	.0228	.0277	.0332	.0393	.0459
5	.0006	.0009	.0014	.0019	.0026	.0035	.0045	.0058	.0074	.0092
6	.0000	.0001	.0001	.0002	.0002	.0003	.0005	.0006	.0009	.0011
7	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001

r \ p	.21	.22	.23	.24	.25	.26	.27	.28	.29	.30
0	.1517	.1370	.1236	.1113	.1001	.0899	.0806	.0722	.0646	.0576
1	.3226	.3092	.2953	.2812	.2670	.2527	.2386	.2247	.2110	.1977
2	.3002	.3052	.3087	.3108	.3115	.3108	.3089	.3058	.3017	.2965
3	.1596	.1722	.1844	.1963	.2076	.2184	.2285	.2379	.2464	.2541
4	.0530	.0607	.0689	.0775	.0865	.0959	.1056	.1156	.1258	.1361
5	.0113	.0137	.0165	.0196	.0231	.0270	.0313	.0360	.0411	.0467
6	.0015	.0019	.0025	.0031	.0038	.0047	.0058	.0070	.0084	.0100
7	.0001	.0002	.0002	.0003	.0004	.0005	.0006	.0008	.0010	.0012
8	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001

r \ p	.31	.32	.33	.34	.35	.36	.37	.38	.39	.40
0	.0514	.0457	.0406	.0360	.0319	.0281	.0248	.0218	.0192	.0168
1	.1847	.1721	.1600	.1484	.1373	.1267	.1166	.1071	.0981	.0896
2	.2904	.2835	.2758	.2675	.2587	.2494	.2397	.2297	.2194	.2090
3	.2609	.2668	.2717	.2756	.2786	.2805	.2815	.2815	.2806	.2787
4	.1465	.1569	.1673	.1775	.1875	.1973	.2067	.2157	.2242	.2322
5	.0527	.0591	.0659	.0732	.0808	.0888	.0971	.1058	.1147	.1239
6	.0118	.0139	.0162	.0188	.0217	.0250	.0285	.0324	.0367	.0413
7	.0015	.0019	.0023	.0028	.0033	.0040	.0048	.0057	.0067	.0079
8	.0001	.0001	.0001	.0002	.0002	.0003	.0004	.0004	.0005	.0007

r \ p	.41	.42	.43	.44	.45	.46	.47	.48	.49	.50
0	.0147	.0128	.0111	.0097	.0084	.0072	.0062	.0053	.0046	.0039
1	.0816	.0742	.0672	.0608	.0548	.0493	.0442	.0395	.0352	.0312
2	.1985	.1880	.1776	.1672	.1569	.1469	.1371	.1275	.1183	.1094
3	.2759	.2723	.2679	.2627	.2568	.2503	.2431	.2355	.2273	.2188
4	.2397	.2465	.2526	.2580	.2627	.2665	.2695	.2717	.2730	.2734
5	.1332	.1428	.1525	.1622	.1719	.1816	.1912	.2006	.2098	.2188
6	.0463	.0517	.0575	.0637	.0703	.0774	.0848	.0926	.1008	.1094
7	.0092	.0107	.0124	.0143	.0164	.0188	.0215	.0244	.0277	.0312
8	.0008	.0010	.0012	.0014	.0017	.0020	.0024	.0028	.0033	.0039

n = 9

r \ p	.01	.02	.03	.04	.05	.06	.07	.08	.09	.10
0	.9135	.8337	.7602	.6925	.6302	.5730	.5204	.4722	.4279	.3874
1	.0830	.1531	.2116	.2597	.2985	.3292	.3525	.3695	.3809	.3874
2	.0034	.0125	.0262	.0433	.0629	.0840	.1061	.1285	.1507	.1722
3	.0001	.0006	.0019	.0042	.0077	.0125	.0186	.0261	.0348	.0446
4	.0000	.0000	.0001	.0003	.0006	.0012	.0021	.0034	.0052	.0074
5	.0000	.0000	.0000	.0000	.0000	.0001	.0002	.0003	.0005	.0008
6	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001

تابع جدول (2)

n = 9

r \ p	.11	.12	.13	.14	.15	.16	.17	.18	.19	.20
0	3504	3165	2855	2573	2316	2082	1869	1676	1501	1342
1	3897	3884	3840	3770	3679	3569	3446	3312	3169	3020
2	1927	2119	2295	2455	2597	2720	2823	2908	2973	3020
3	0556	0674	0800	0833	1069	1209	1349	1489	1627	1762
4	0103	0138	0178	0228	0283	0345	0415	0490	0573	0661
5	0013	0019	0027	0037	0050	0066	0085	0108	0134	0165
6	0001	0002	0003	0004	0006	0008	0012	0016	0021	0028
7	0000	0000	0000	0000	0000	0001	0001	0001	0002	0003

r \ p	.21	.22	.23	.24	.25	.26	.27	.28	.29	.30
0	1199	1069	0952	0846	0751	0665	0589	0520	0458	0404
1	2867	2713	2558	2404	2253	2104	1960	1820	1685	1556
2	3049	3061	3056	3037	3003	2957	2899	2831	2754	2668
3	1891	2014	2130	2238	2336	2424	2502	2569	2624	2668
4	0754	0852	0954	1060	1168	1278	1388	1499	1608	1715
5	0200	0240	0285	0335	0389	0449	0513	0583	0657	0735
6	0036	0045	0057	0070	0087	0105	0127	0151	0179	0210
7	0004	0005	0007	0010	0012	0016	0020	0025	0031	0039
8	0000	0000	0001	0001	0001	0001	0002	0002	0003	0004

r \ p	.31	.32	.33	.34	.35	.36	.37	.38	.39	.40
0	0355	0311	0272	0238	0207	0180	0156	0135	0117	0101
1	1433	1317	1206	1103	1004	0912	0826	0747	0673	0605
2	2576	2478	2376	2270	2162	2052	1941	1831	1721	1612
3	2701	2721	2731	2729	2716	2693	2660	2618	2567	2508
4	1820	1921	2017	2109	2194	2272	2344	2407	2462	2508
5	0818	0904	0994	1086	1181	1278	1376	1475	1574	1672
6	0245	0284	0326	0373	0424	0479	0539	0603	0671	0743
7	0047	0057	0069	0082	0098	0116	0136	0158	0184	0212
8	0005	0007	0008	0011	0013	0016	0020	0024	0029	0035
9	0000	0000	0000	0001	0001	0001	0001	0002	0002	0003

r \ p	.41	.42	.43	.44	.45	.46	.47	.48	.49	.50
0	0087	0074	0064	0054	0046	0039	0033	0028	0023	0020
1	0542	0484	0431	0383	0339	0299	0263	0231	0202	0176
2	1506	1402	1301	1204	1110	1020	0934	0853	0776	0703
3	2442	2369	2291	2207	2119	2027	1933	1837	1739	1641
4	2545	2573	2592	2601	2600	2590	2571	2543	2506	2461
5	1769	1863	1955	2044	2128	2207	2280	2347	2408	2461
6	0819	0900	0983	1070	1160	1253	1348	1445	1542	1641
7	0244	0279	0318	0360	0407	0458	0512	0571	0635	0703
8	0042	0051	0060	0071	0083	0097	0114	0132	0153	0176
9	0003	0004	0005	0006	0008	0009	0011	0014	0016	0020

n = 10

r \ p	.01	.02	.03	.04	.05	.06	.07	.08	.09	.10
0	9044	8171	7374	6648	5987	5386	4840	4344	3894	3487
1	0914	1667	2281	2770	3151	3438	3643	3777	3851	3874
2	0042	0153	0317	0519	0746	0988	1234	1478	1714	1937
3	0001	0008	0026	0058	0105	0168	0248	0343	0452	0574
4	0000	0000	0001	0004	0010	0019	0033	0052	0078	0112
5	0000	0000	0000	0000	0001	0001	0003	0005	0009	0015
6	0000	0000	0000	0000	0000	0000	0000	0000	0001	0001

r \ p	.11	.12	.13	.14	.15	.16	.17	.18	.19	.20
0	3118	2785	2484	2213	1969	1749	1552	1374	1216	1074
1	3854	3798	3712	3603	3474	3331	3178	3017	2852	2684
2	2143	2330	2496	2639	2759	2856	2929	2980	3010	3020
3	0706	0847	0995	1146	1298	1450	1600	1745	1883	2013
4	0153	0202	0260	0326	0401	0483	0573	0670	0773	0881
5	0023	0033	0047	0064	0085	0111	0141	0177	0218	0264
6	0002	0004	0006	0009	0012	0018	0024	0032	0043	0055
7	0000	0000	0000	0001	0001	0002	0003	0004	0006	0008
8	0000	0000	0000	0000	0000	0000	0000	0000	0001	0001

تابع جدول (2)

n = 10

r \ p	.21	.22	.23	.24	.25	.26	.27	.28	.29	.30
0	.0947	.0834	.0733	.0643	.0563	.0492	.0430	.0374	.0326	.0282
1	.2517	.2351	.2188	.2030	.1877	.1730	.1590	.1456	.1330	.1211
2	.3011	.2984	.2942	.2885	.2816	.2735	.2646	.2548	.2444	.2335
3	.2134	.2244	.2343	.2429	.2503	.2563	.2609	.2642	.2662	.2668
4	.0993	.1108	.1225	.1343	.1460	.1576	.1689	.1798	.1903	.2001
5	.0317	.0375	.0439	.0509	.0584	.0664	.0750	.0839	.0933	.1029
6	.0070	.0088	.0109	.0134	.0162	.0195	.0231	.0272	.0317	.0368
7	.0011	.0014	.0019	.0024	.0031	.0039	.0049	.0060	.0074	.0090
8	.0001	.0002	.0002	.0003	.0004	.0005	.0007	.0009	.0011	.0014
9	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0001	.0001

r \ p	.31	.32	.33	.34	.35	.36	.37	.38	.39	.40
0	.0245	.0211	.0182	.0157	.0135	.0115	.0098	.0084	.0071	.0060
1	.1099	.0995	.0898	.0808	.0725	.0649	.0578	.0514	.0456	.0403
2	.2222	.2107	.1990	.1873	.1757	.1642	.1529	.1419	.1312	.1209
3	.2662	.2644	.2614	.2573	.2522	.2462	.2394	.2319	.2237	.2150
4	.2093	.2177	.2253	.2320	.2377	.2424	.2461	.2487	.2503	.2508
5	.1128	.1229	.1332	.1434	.1536	.1636	.1734	.1829	.1920	.2007
6	.0422	.0482	.0547	.0616	.0689	.0767	.0849	.0934	.1023	.1115
7	.0108	.0130	.0154	.0181	.0212	.0247	.0285	.0327	.0374	.0425
8	.0018	.0023	.0028	.0035	.0043	.0052	.0063	.0075	.0090	.0106
9	.0002	.0002	.0003	.0004	.0005	.0006	.0008	.0010	.0013	.0016
10	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0001

r \ p	.41	.42	.43	.44	.45	.46	.47	.48	.49	.50
0	.0051	.0043	.0036	.0030	.0025	.0021	.0017	.0014	.0012	.0010
1	.0355	.0312	.0273	.0238	.0207	.0180	.0155	.0133	.0114	.0098
2	.1111	.1017	.0927	.0843	.0763	.0688	.0619	.0554	.0494	.0439
3	.2058	.1963	.1865	.1765	.1665	.1564	.1464	.1364	.1267	.1172
4	.2503	.2488	.2462	.2427	.2384	.2331	.2271	.2204	.2130	.2051
5	.2087	.2162	.2229	.2289	.2340	.2383	.2417	.2441	.2456	.2461
6	.1209	.1304	.1401	.1499	.1596	.1692	.1786	.1878	.1966	.2051
7	.0480	.0540	.0604	.0673	.0746	.0824	.0905	.0991	.1080	.1172
8	.0125	.0147	.0171	.0198	.0229	.0263	.0301	.0343	.0389	.0439
9	.0019	.0024	.0029	.0035	.0042	.0050	.0059	.0070	.0083	.0098
10	.0001	.0002	.0002	.0003	.0003	.0004	.0005	.0006	.0008	.0010

n = 11

r \ p	.01	.02	.03	.04	.05	.06	.07	.08	.09	.10
0	.8953	.8007	.7153	.6382	.5688	.5063	.4501	.3996	.3544	.3138
1	.0995	.1798	.2433	.2925	.3293	.3555	.3727	.3823	.3855	.3835
2	.0050	.0183	.0376	.0609	.0867	.1135	.1403	.1662	.1906	.2131
3	.0002	.0011	.0035	.0076	.0137	.0217	.0317	.0434	.0560	.0710
4	.0000	.0000	.0002	.0006	.0014	.0028	.0048	.0075	.0112	.0158
5	.0000	.0000	.0000	.0000	.0001	.0002	.0005	.0009	.0015	.0025
6	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0002	.0003

r \ p	.11	.12	.13	.14	.15	.16	.17	.18	.19	.20
0	.2775	.2451	.2161	.1903	.1673	.1468	.1288	.1127	.0985	.0859
1	.3773	.3676	.3552	.3408	.3248	.3078	.2901	.2721	.2541	.2362
2	.2332	.2507	.2654	.2774	.2866	.2932	.2971	.2987	.2980	.2953
3	.0865	.1025	.1190	.1355	.1517	.1675	.1826	.1967	.2097	.2215
4	.0214	.0280	.0356	.0441	.0536	.0638	.0748	.0864	.0984	.1107
5	.0037	.0053	.0074	.0101	.0132	.0170	.0214	.0265	.0323	.0388
6	.0005	.0007	.0011	.0016	.0023	.0032	.0044	.0058	.0076	.0097
7	.0000	.0001	.0001	.0002	.0003	.0004	.0006	.0008	.0013	.0017
8	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0001	.0002

تابع جدول (2)

n = 11

r \ p	.21	.22	.23	.24	.25	.26	.27	.28	.29	.30
0	.0748	.0650	.0564	.0489	.0422	.0364	.0314	.0270	.0231	.0198
1	.2187	.2017	.1854	.1697	.1549	.1408	.1278	.1153	.1038	.0932
2	.2907	.2845	.2768	.2680	.2581	.2474	.2360	.2242	.2121	.1998
3	.2318	.2407	.2481	.2539	.2581	.2608	.2619	.2616	.2599	.2568
4	.1232	.1358	.1482	.1603	.1721	.1832	.1937	.2035	.2123	.2201
5	.0459	.0536	.0620	.0709	.0803	.0901	.1003	.1108	.1214	.1321
6	.0122	.0151	.0185	.0224	.0268	.0317	.0371	.0431	.0496	.0566
7	.0023	.0030	.0039	.0050	.0064	.0079	.0098	.0120	.0145	.0173
8	.0003	.0004	.0006	.0008	.0011	.0014	.0018	.0023	.0030	.0037
9	.0000	.0000	.0001	.0001	.0001	.0002	.0002	.0003	.0004	.0005

r \ p	.31	.32	.33	.34	.35	.36	.37	.38	.39	.40
0	.0169	.0144	.0122	.0104	.0088	.0074	.0062	.0052	.0044	.0036
1	.0834	.0744	.0662	.0587	.0518	.0457	.0401	.0351	.0306	.0266
2	.1874	.1751	.1630	.1511	.1395	.1284	.1177	.1075	.0978	.0887
3	.2526	.2472	.2408	.2335	.2254	.2167	.2074	.1977	.1876	.1774
4	.2269	.2326	.2372	.2406	.2428	.2438	.2436	.2423	.2399	.2365
5	.1427	.1533	.1636	.1735	.1830	.1920	.2003	.2079	.2148	.2207
6	.0641	.0721	.0806	.0894	.0985	.1080	.1176	.1274	.1373	.1471
7	.0206	.0242	.0283	.0329	.0379	.0434	.0494	.0558	.0627	.0701
8	.0046	.0057	.0070	.0085	.0102	.0122	.0145	.0171	.0200	.0234
9	.0007	.0009	.0011	.0015	.0018	.0023	.0028	.0035	.0043	.0052
10	.0001	.0001	.0001	.0001	.0002	.0003	.0003	.0004	.0005	.0007

r \ p	.41	.42	.43	.44	.45	.46	.47	.48	.49	.50
0	.0030	.0025	.0021	.0017	.0014	.0011	.0009	.0008	.0006	.0005
1	.0231	.0199	.0171	.0147	.0125	.0107	.0090	.0078	.0064	.0054
2	.0801	.0721	.0646	.0577	.0513	.0454	.0401	.0352	.0308	.0269
3	.1670	.1568	.1462	.1359	.1259	.1161	.1067	.0976	.0888	.0806
4	.2321	.2267	.2206	.2136	.2060	.1978	.1892	.1801	.1707	.1611
5	.2258	.2298	.2329	.2350	.2360	.2359	.2348	.2327	.2296	.2256
6	.1569	.1664	.1757	.1846	.1931	.2010	.2083	.2148	.2206	.2256
7	.0779	.0861	.0947	.1036	.1128	.1223	.1318	.1416	.1514	.1611
8	.0271	.0312	.0357	.0407	.0462	.0521	.0585	.0654	.0727	.0806
9	.0063	.0075	.0090	.0107	.0126	.0148	.0173	.0201	.0233	.0269
10	.0009	.0011	.0014	.0017	.0021	.0025	.0031	.0037	.0045	.0054
11	.0001	.0001	.0001	.0001	.0002	.0002	.0003	.0003	.0004	.0005

n = 12

r \ p	.01	.02	.03	.04	.05	.06	.07	.08	.09	.10
0	.8864	.7847	.6938	.6127	.5404	.4759	.4186	.3677	.3225	.2824
1	.1074	.1922	.2575	.3064	.3413	.3645	.3781	.3837	.3827	.3766
2	.0060	.0216	.0438	.0702	.0988	.1280	.1565	.1835	.2082	.2301
3	.0002	.0015	.0045	.0098	.0173	.0272	.0393	.0532	.0688	.0852
4	.0000	.0001	.0003	.0009	.0021	.0039	.0067	.0104	.0153	.0213
5	.0000	.0000	.0000	.0001	.0002	.0004	.0008	.0014	.0024	.0038
6	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0003	.0005

r \ p	.11	.12	.13	.14	.15	.16	.17	.18	.19	.20
0	.2470	.2157	.1880	.1637	.1422	.1234	.1069	.0924	.0798	.0687
1	.3663	.3529	.3372	.3197	.3012	.2821	.2627	.2434	.2245	.2062
2	.2490	.2647	.2771	.2863	.2924	.2955	.2960	.2939	.2897	.2835
3	.1026	.1203	.1380	.1553	.1720	.1876	.2021	.2151	.2265	.2362
4	.0285	.0369	.0464	.0569	.0683	.0804	.0931	.1062	.1195	.1329
5	.0056	.0081	.0111	.0148	.0193	.0245	.0305	.0373	.0449	.0532
6	.0008	.0013	.0019	.0028	.0040	.0054	.0073	.0096	.0132	.0155
7	.0001	.0001	.0002	.0004	.0006	.0008	.0013	.0018	.0025	.0033
8	.0000	.0000	.0000	.0000	.0001	.0001	.0002	.0002	.0004	.0005
9	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001

تابع جدول (2)

n = 12

r \ p	.21	.22	.23	.24	.25	.26	.27	.28	.29	.30
0	.0591	.0507	.0434	.0371	.0317	.0270	.0229	.0194	.0164	.0138
1	.1885	.1717	.1557	.1407	.1267	.1137	.1016	.0906	.0804	.0712
2	.2756	.2663	.2558	.2444	.2323	.2197	.2068	.1937	.1807	.1678
3	.2442	.2503	.2547	.2573	.2581	.2573	.2549	.2511	.2460	.2397
4	.1460	.1589	.1712	.1828	.1936	.2034	.2122	.2197	.2261	.2311
5	.0621	.0717	.0818	.0924	.1032	.1143	.1255	.1367	.1477	.1585
6	.0193	.0236	.0285	.0340	.0401	.0469	.0542	.0620	.0704	.0792
7	.0044	.0057	.0073	.0092	.0115	.0141	.0172	.0207	.0246	.0291
8	.0007	.0010	.0014	.0018	.0024	.0031	.0040	.0050	.0063	.0078
9	.0001	.0001	.0002	.0003	.0004	.0005	.0007	.0009	.0011	.0015
10	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0001	.0001	.0002

r \ p	.31	.32	.33	.34	.35	.36	.37	.38	.39	.40
0	.0116	.0098	.0082	.0068	.0057	.0047	.0039	.0032	.0027	.0022
1	.0628	.0552	.0484	.0422	.0368	.0319	.0276	.0237	.0204	.0174
2	.1552	.1429	.1310	.1197	.1088	.0986	.0890	.0800	.0716	.0639
3	.2324	.2241	.2151	.2055	.1954	.1849	.1742	.1634	.1526	.1419
4	.2349	.2373	.2384	.2382	.2367	.2340	.2302	.2254	.2195	.2128
5	.1688	.1787	.1879	.1963	.2039	.2106	.2163	.2210	.2246	.2270
6	.0885	.0981	.1079	.1180	.1281	.1382	.1482	.1580	.1675	.1766
7	.0341	.0396	.0456	.0521	.0591	.0666	.0746	.0830	.0918	.1009
8	.0096	.0116	.0140	.0168	.0199	.0234	.0274	.0318	.0367	.0420
9	.0019	.0024	.0031	.0038	.0048	.0059	.0071	.0087	.0104	.0125
10	.0003	.0003	.0005	.0006	.0008	.0010	.0013	.0016	.0020	.0025
11	.0000	.0000	.0000	.0001	.0001	.0001	.0001	.0002	.0002	.0003

r \ p	.41	.42	.43	.44	.45	.46	.47	.48	.49	.50
0	.0018	.0014	.0012	.0010	.0008	.0006	.0005	.0004	.0003	.0002
1	.0148	.0126	.0106	.0090	.0075	.0063	.0052	.0043	.0036	.0029
2	.0567	.0502	.0442	.0388	.0339	.0294	.0255	.0220	.0189	.0161
3	.1314	.1211	.1111	.1015	.0923	.0836	.0754	.0676	.0604	.0537
4	.2054	.1973	.1886	.1794	.1700	.1602	.1504	.1405	.1306	.1208
5	.2284	.2285	.2276	.2256	.2225	.2184	.2134	.2075	.2008	.1934
6	.1851	.1931	.2003	.2068	.2124	.2171	.2208	.2234	.2250	.2256
7	.1103	.1198	.1295	.1393	.1489	.1585	.1678	.1768	.1853	.1934
8	.0479	.0542	.0611	.0684	.0762	.0844	.0930	.1020	.1113	.1208
9	.0148	.0175	.0205	.0239	.0277	.0319	.0367	.0418	.0475	.0537
10	.0031	.0038	.0046	.0056	.0068	.0082	.0098	.0116	.0137	.0161
11	.0004	.0005	.0006	.0008	.0010	.0013	.0016	.0019	.0024	.0029
12	.0000	.0000	.0000	.0001	.0001	.0001	.0001	.0001	.0002	.0002

n = 13

r \ p	.01	.02	.03	.04	.05	.06	.07	.08	.09	.10
0	.8775	.7690	.6730	.5882	.5133	.4474	.3893	.3383	.2935	.2542
1	.1152	.2040	.2706	.3186	.3512	.3712	.3809	.3824	.3773	.3672
2	.0070	.0250	.0502	.0797	.1109	.1422	.1720	.1995	.2239	.2448
3	.0003	.0019	.0057	.0122	.0214	.0333	.0475	.0636	.0812	.0997
4	.0000	.0001	.0004	.0013	.0028	.0053	.0089	.0138	.0201	.0277
5	.0000	.0000	.0000	.0001	.0003	.0006	.0012	.0022	.0036	.0055
6	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0003	.0005	.0008
7	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001

r \ p	.11	.12	.13	.14	.15	.16	.17	.18	.19	.20
0	.2198	.1898	.1636	.1408	.1209	.1037	.0887	.0758	.0646	.0550
1	.3532	.3364	.3178	.2979	.2774	.2567	.2362	.2163	.1970	.1787
2	.2619	.2753	.2849	.2910	.2937	.2934	.2903	.2848	.2773	.2680
3	.1187	.1376	.1561	.1737	.1900	.2049	.2180	.2293	.2385	.2457
4	.0367	.0469	.0583	.0707	.0838	.0976	.1116	.1258	.1399	.1535
5	.0082	.0115	.0157	.0207	.0266	.0335	.0412	.0497	.0591	.0691
6	.0013	.0021	.0031	.0045	.0063	.0085	.0112	.0145	.0185	.0230
7	.0002	.0003	.0005	.0007	.0011	.0016	.0023	.0032	.0043	.0058
8	.0000	.0000	.0001	.0001	.0001	.0002	.0004	.0005	.0008	.0011
9	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0001

تابع جدول (2)

n = 13

r \ p	.21	.22	.23	.24	.25	.26	.27	.28	.29	.30
0	.0467	.0396	.0334	.0282	.0238	.0200	.0167	.0140	.0117	.0097
1	.1613	.1450	.1299	.1159	.1029	.0911	.0804	.0706	.0619	.0540
2	.2573	.2455	.2328	.2195	.2059	.1921	.1784	.1648	.1516	.1388
3	.2508	.2539	.2550	.2542	.2517	.2475	.2419	.2351	.2271	.2181
4	.1667	.1790	.1904	.2007	.2097	.2174	.2237	.2285	.2319	.2337
5	.0797	.0909	.1024	.1141	.1258	.1375	.1489	.1600	.1705	.1803
6	.0283	.0342	.0408	.0480	.0559	.0644	.0734	.0829	.0928	.1030
7	.0075	.0096	.0122	.0152	.0186	.0226	.0272	.0323	.0379	.0442
8	.0015	.0020	.0027	.0036	.0047	.0060	.0075	.0094	.0116	.0142
9	.0002	.0003	.0005	.0006	.0009	.0012	.0015	.0020	.0026	.0034
10	.0000	.0000	.0001	.0001	.0001	.0002	.0002	.0003	.0004	.0006
11	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001

r \ p	.31	.32	.33	.34	.35	.36	.37	.38	.39	.40
0	.0080	.0066	.0055	.0045	.0037	.0030	.0025	.0020	.0016	.0013
1	.0469	.0407	.0351	.0302	.0259	.0221	.0188	.0159	.0135	.0113
2	.1265	.1148	.1037	.0933	.0836	.0746	.0663	.0586	.0516	.0453
3	.2084	.1981	.1874	.1763	.1651	.1538	.1427	.1317	.1210	.1107
4	.2341	.2331	.2307	.2270	.2222	.2163	.2095	.2018	.1934	.1845
5	.1893	.1974	.2045	.2105	.2154	.2190	.2215	.2227	.2226	.2214
6	.1134	.1239	.1343	.1446	.1546	.1643	.1734	.1820	.1898	.1968
7	.0509	.0583	.0662	.0745	.0833	.0924	.1019	.1115	.1213	.1312
8	.0172	.0206	.0244	.0288	.0336	.0390	.0448	.0513	.0582	.0656
9	.0043	.0054	.0067	.0082	.0101	.0122	.0146	.0175	.0207	.0243
10	.0008	.0010	.0013	.0017	.0022	.0027	.0034	.0043	.0053	.0065
11	.0001	.0001	.0002	.0002	.0003	.0004	.0006	.0007	.0009	.0012
12	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0001	.0001

r \ p	.41	.42	.43	.44	.45	.46	.47	.48	.49	.50
0	.0010	.0008	.0007	.0005	.0004	.0003	.0003	.0002	.0002	.0001
1	.0095	.0079	.0068	.0054	.0045	.0037	.0030	.0024	.0020	.0016
2	.0395	.0344	.0298	.0256	.0220	.0188	.0160	.0135	.0114	.0095
3	.1007	.0913	.0823	.0739	.0660	.0587	.0519	.0457	.0401	.0349
4	.1750	.1653	.1553	.1451	.1350	.1250	.1151	.1055	.0962	.0873
5	.2189	.2154	.2108	.2053	.1989	.1917	.1838	.1753	.1664	.1571
6	.2029	.2080	.2121	.2151	.2169	.2177	.2173	.2158	.2131	.2095
7	.1410	.1506	.1600	.1690	.1775	.1854	.1927	.1992	.2048	.2095
8	.0735	.0818	.0905	.0996	.1089	.1185	.1282	.1379	.1476	.1571
9	.0284	.0329	.0379	.0435	.0495	.0561	.0631	.0707	.0788	.0873
10	.0079	.0095	.0114	.0137	.0162	.0191	.0224	.0261	.0303	.0349
11	.0015	.0019	.0024	.0029	.0038	.0044	.0054	.0066	.0079	.0095
12	.0002	.0002	.0003	.0004	.0005	.0006	.0008	.0010	.0013	.0016
13	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0001	.0001

n = 14

r \ p	.01	.02	.03	.04	.05	.06	.07	.08	.09	.10
0	.8687	.7536	.6528	.5647	.4877	.4205	.3620	.3112	.2670	.2288
1	.1229	.2153	.2827	.3294	.3593	.3758	.3815	.3788	.3698	.3559
2	.0081	.0286	.0568	.0892	.1229	.1559	.1867	.2141	.2377	.2570
3	.0003	.0023	.0070	.0149	.0259	.0398	.0562	.0745	.0940	.1142
4	.0000	.0001	.0006	.0017	.0037	.0070	.0116	.0178	.0256	.0349
5	.0000	.0000	.0000	.0001	.0004	.0009	.0018	.0031	.0051	.0078
6	.0000	.0000	.0000	.0000	.0000	.0001	.0002	.0004	.0008	.0013
7	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0002

r \ p	.11	.12	.13	.14	.15	.16	.17	.18	.19	.20
0	.1956	.1670	.1423	.1211	.1028	.0871	.0736	.0621	.0523	.0440
1	.3385	.3188	.2977	.2759	.2539	.2322	.2112	.1910	.1719	.1539
2	.2720	.2826	.2892	.2919	.2912	.2875	.2811	.2725	.2620	.2501
3	.1345	.1542	.1728	.1901	.2056	.2190	.2303	.2393	.2459	.2501
4	.0457	.0578	.0710	.0851	.0998	.1147	.1297	.1444	.1586	.1720
5	.0113	.0158	.0212	.0277	.0352	.0437	.0531	.0634	.0744	.0860
6	.0021	.0032	.0048	.0068	.0093	.0125	.0163	.0209	.0262	.0322
7	.0003	.0005	.0008	.0013	.0019	.0027	.0038	.0052	.0070	.0092
8	.0000	.0001	.0001	.0002	.0003	.0005	.0007	.0010	.0014	.0020
9	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0001	.0002	.0003

تابع جدول (2)

n = 14

r \ p	.21	.22	.23	.24	.25	.26	.27	.28	.29	.30
0	.0369	.0309	.0258	.0214	.0178	.0148	.0122	.0101	.0083	.0068
1	.1372	.1218	.1077	.0948	.0832	.0726	.0632	.0548	.0473	.0407
2	.2371	.2234	.2091	.1946	.1802	.1659	.1519	.1385	.1256	.1134
3	.2521	.2520	.2499	.2459	.2402	.2331	.2248	.2154	.2052	.1943
4	.1843	.1955	.2052	.2135	.2202	.2252	.2286	.2304	.2305	.2290
5	.0980	.1103	.1226	.1348	.1468	.1583	.1691	.1792	.1883	.1963
6	.0391	.0466	.0549	.0639	.0734	.0834	.0938	.1045	.1153	.1262
7	.0119	.0150	.0188	.0231	.0280	.0335	.0397	.0464	.0538	.0618
8	.0028	.0037	.0049	.0064	.0082	.0103	.0128	.0158	.0192	.0232
9	.0005	.0007	.0010	.0013	.0018	.0024	.0032	.0041	.0052	.0066
10	.0001	.0001	.0001	.0002	.0003	.0004	.0006	.0008	.0011	.0014
11	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0001	.0002	.0002

r \ p	.31	.32	.33	.34	.35	.36	.37	.38	.39	.40
0	.0055	.0045	.0037	.0030	.0024	.0019	.0016	.0012	.0010	.0008
1	.0349	.0298	.0253	.0215	.0181	.0152	.0126	.0106	.0088	.0073
2	.1018	.0911	.0811	.0719	.0634	.0557	.0487	.0424	.0367	.0317
3	.1830	.1715	.1598	.1481	.1366	.1253	.1144	.1039	.0940	.0845
4	.2261	.2219	.2164	.2098	.2022	.1938	.1848	.1752	.1652	.1549
5	.2032	.2088	.2132	.2161	.2178	.2181	.2170	.2147	.2112	.2066
6	.1369	.1474	.1575	.1670	.1759	.1840	.1912	.1974	.2026	.2066
7	.0703	.0793	.0886	.0983	.1082	.1183	.1283	.1383	.1480	.1574
8	.0276	.0326	.0382	.0443	.0510	.0582	.0659	.0742	.0828	.0918
9	.0083	.0102	.0125	.0152	.0183	.0218	.0258	.0303	.0353	.0408
10	.0019	.0024	.0031	.0039	.0049	.0061	.0076	.0093	.0113	.0136
11	.0003	.0004	.0006	.0007	.0010	.0013	.0016	.0021	.0026	.0033
12	.0000	.0000	.0001	.0001	.0001	.0002	.0002	.0003	.0004	.0005
13	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001

r \ p	.41	.42	.43	.44	.45	.46	.47	.48	.49	.50
0	.0006	.0005	.0004	.0003	.0002	.0002	.0001	.0001	.0001	.0001
1	.0060	.0049	.0040	.0033	.0027	.0021	.0017	.0014	.0011	.0009
2	.0272	.0233	.0198	.0168	.0141	.0118	.0099	.0082	.0068	.0056
3	.0757	.0674	.0597	.0527	.0462	.0403	.0350	.0303	.0260	.0222
4	.1446	.1342	.1239	.1138	.1040	.0945	.0854	.0768	.0687	.0611
5	.2009	.1943	.1869	.1788	.1701	.1610	.1515	.1418	.1320	.1222
6	.2094	.2111	.2115	.2108	.2088	.2057	.2015	.1963	.1902	.1833
7	.1663	.1747	.1824	.1892	.1952	.2003	.2043	.2071	.2089	.2095
8	.1011	.1107	.1204	.1301	.1398	.1493	.1585	.1673	.1756	.1833
9	.0469	.0534	.0605	.0682	.0762	.0848	.0937	.1030	.1125	.1222
10	.0163	.0193	.0228	.0268	.0312	.0361	.0415	.0475	.0540	.0611
11	.0041	.0051	.0063	.0076	.0093	.0112	.0134	.0160	.0189	.0222
12	.0007	.0009	.0012	.0015	.0019	.0024	.0030	.0037	.0045	.0056
13	.0001	.0001	.0001	.0002	.0002	.0003	.0004	.0005	.0007	.0009
14	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001

n = 15

r \ p	.01	.02	.03	.04	.05	.06	.07	.08	.09	.10
0	.8601	.7386	.6333	.5421	.4633	.3953	.3367	.2863	.2430	.2059
1	.1303	.2261	.2938	.3388	.3658	.3785	.3801	.3734	.3605	.3432
2	.0092	.0323	.0636	.0988	.1348	.1691	.2003	.2273	.2496	.2669
3	.0004	.0029	.0085	.0178	.0307	.0468	.0653	.0857	.1070	.1285
4	.0000	.0002	.0008	.0022	.0049	.0090	.0148	.0223	.0317	.0428
5	.0000	.0000	.0001	.0002	.0006	.0013	.0024	.0043	.0068	.0105
6	.0000	.0000	.0000	.0000	.0000	.0001	.0003	.0006	.0011	.0019
7	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0003

تابع جدول (2)

n = 15

r \ p	.11	.12	.13	.14	.15	.16	.17	.18	.19	.20
0	.1741	.1470	.1238	.1041	.0874	.0731	.0611	.0510	.0424	.0352
1	.3228	.3006	.2775	.2542	.2312	.2090	.1878	.1678	.1492	.1319
2	.2793	.2870	.2903	.2897	.2856	.2787	.2692	.2578	.2449	.2309
3	.1496	.1696	.1880	.2044	.2184	.2300	.2389	.2452	.2459	.2501
4	.0555	.0694	.0843	.0998	.1156	.1314	.1468	.1615	.1752	.1876
5	.0151	.0208	.0277	.0357	.0449	.0551	.0662	.0780	.0904	.1032
6	.0031	.0047	.0069	.0097	.0132	.0175	.0226	.0285	.0353	.0430
7	.0005	.0008	.0013	.0020	.0030	.0043	.0059	.0081	.0107	.0138
8	.0001	.0001	.0002	.0003	.0005	.0008	.0012	.0018	.0025	.0035
9	.0000	.0000	.0000	.0000	.0001	.0001	.0002	.0003	.0005	.0007
10	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001

r \ p	.21	.22	.23	.24	.25	.26	.27	.28	.29	.30
0	.0291	.0241	.0198	.0163	.0134	.0109	.0089	.0072	.0059	.0047
1	.1162	.1018	.0889	.0772	.0668	.0576	.0494	.0423	.0360	.0305
2	.2162	.2010	.1858	.1707	.1559	.1416	.1280	.1150	.1029	.0916
3	.2490	.2457	.2405	.2336	.2252	.2156	.2051	.1939	.1821	.1700
4	.1986	.2079	.2155	.2213	.2252	.2273	.2276	.2262	.2231	.2186
5	.1161	.1290	.1416	.1537	.1651	.1757	.1852	.1935	.2005	.2061
6	.0514	.0606	.0705	.0809	.0917	.1029	.1142	.1254	.1365	.1472
7	.0176	.0220	.0271	.0329	.0393	.0465	.0543	.0627	.0717	.0811
8	.0047	.0062	.0081	.0104	.0131	.0163	.0201	.0244	.0293	.0348
9	.0010	.0014	.0019	.0025	.0034	.0045	.0058	.0074	.0093	.0116
10	.0002	.0002	.0003	.0005	.0007	.0009	.0013	.0017	.0023	.0030
11	.0000	.0000	.0000	.0001	.0001	.0002	.0002	.0003	.0004	.0006
12	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001

r \ p	.31	.32	.33	.34	.35	.36	.37	.38	.39	.40
0	.0038	.0031	.0025	.0020	.0016	.0012	.0010	.0008	.0006	.0005
1	.0258	.0217	.0182	.0152	.0126	.0104	.0086	.0071	.0058	.0047
2	.0811	.0715	.0627	.0547	.0476	.0411	.0354	.0303	.0259	.0219
3	.1579	.1457	.1338	.1222	.1110	.1002	.0901	.0805	.0716	.0634
4	.2128	.2057	.1977	.1888	.1792	.1692	.1587	.1481	.1374	.1268
5	.2110	.2130	.2142	.2140	.2123	.2093	.2051	.1997	.1933	.1859
6	.1575	.1671	.1759	.1837	.1906	.1963	.2008	.2040	.2059	.2066
7	.0910	.1011	.1114	.1217	.1319	.1419	.1516	.1608	.1693	.1771
8	.0409	.0476	.0549	.0627	.0710	.0798	.0890	.0985	.1082	.1181
9	.0143	.0174	.0210	.0251	.0298	.0349	.0407	.0470	.0538	.0612
10	.0038	.0049	.0062	.0078	.0096	.0118	.0143	.0173	.0206	.0245
11	.0008	.0011	.0014	.0018	.0024	.0030	.0038	.0048	.0060	.0074
12	.0001	.0002	.0002	.0003	.0004	.0006	.0007	.0010	.0013	.0016
13	.0000	.0000	.0000	.0000	.0001	.0001	.0001	.0001	.0002	.0003

r \ p	.41	.42	.43	.44	.45	.46	.47	.48	.49	.50
0	.0004	.0003	.0002	.0002	.0001	.0001	.0001	.0001	.0000	.0000
1	.0038	.0031	.0025	.0020	.0016	.0012	.0010	.0008	.0006	.0005
2	.0185	.0156	.0130	.0108	.0090	.0074	.0060	.0049	.0040	.0032
3	.0558	.0489	.0426	.0369	.0318	.0272	.0232	.0197	.0164	.0139
4	.1163	.1061	.0963	.0869	.0780	.0698	.0617	.0545	.0478	.0417
5	.1778	.1691	.1598	.1502	.1404	.1304	.1204	.1106	.1010	.0916
6	.2060	.2041	.2010	.1967	.1914	.1851	.1780	.1702	.1617	.1527
7	.1840	.1900	.1949	.1987	.2013	.2028	.2030	.2020	.1997	.1964
8	.1279	.1376	.1470	.1561	.1647	.1727	.1800	.1864	.1919	.1964
9	.0691	.0775	.0863	.0954	.1048	.1144	.1241	.1338	.1434	.1527
10	.0288	.0337	.0390	.0450	.0515	.0585	.0661	.0741	.0827	.0916
11	.0091	.0111	.0134	.0161	.0191	.0226	.0266	.0311	.0361	.0417
12	.0021	.0027	.0034	.0042	.0052	.0064	.0079	.0096	.0138	.0139
13	.0003	.0004	.0006	.0008	.0010	.0013	.0018	.0020	.0026	.0032
14	.0000	.0000	.0001	.0001	.0001	.0002	.0002	.0003	.0004	.0005

تابع جدول (2)

n = 16

r \ p	.01	.02	.03	.04	.05	.06	.07	.08	.09	.10
0	.8515	.7238	.6143	.5204	.4401	.3716	.3131	.2634	.2211	.1853
1	.1376	.2363	.3040	.3469	.3706	.3795	.3771	.3665	.3499	.3294
2	.0104	.0362	.0705	.1084	.1463	.1817	.2129	.2390	.2596	.2745
3	.0005	.0034	.0102	.0211	.0359	.0541	.0748	.0970	.1198	.1423
4	.0000	.0002	.0010	.0029	.0061	.0112	.0183	.0274	.0385	.0514
5	.0000	.0000	.0001	.0003	.0008	.0017	.0033	.0057	.0091	.0137
6	.0000	.0000	.0000	.0000	.0001	.0002	.0005	.0009	.0017	.0028
7	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0002	.0004
8	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001

r \ p	.11	.12	.13	.14	.15	.16	.17	.18	.19	.20
0	.1550	.1293	.1077	.0895	.0743	.0614	.0507	.0418	.0343	.0281
1	.3065	.2822	.2575	.2332	.2097	.1873	.1662	.1468	.1289	.1126
2	.2841	.2886	.2886	.2847	.2775	.2675	.2554	.2416	.2267	.2111
3	.1638	.1837	.2013	.2163	.2285	.2378	.2441	.2475	.2482	.2463
4	.0658	.0814	.0977	.1144	.1311	.1472	.1625	.1766	.1892	.2001
5	.0195	.0266	.0351	.0447	.0555	.0673	.0799	.0930	.1065	.1201
6	.0044	.0067	.0096	.0133	.0180	.0235	.0300	.0374	.0458	.0550
7	.0008	.0013	.0020	.0031	.0045	.0064	.0088	.0117	.0153	.0197
8	.0001	.0002	.0003	.0006	.0009	.0014	.0020	.0029	.0041	.0055
9	.0000	.0000	.0000	.0001	.0001	.0002	.0004	.0006	.0008	.0012
10	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0001	.0002

r \ p	.21	.22	.23	.24	.25	.26	.27	.28	.29	.30
0	.0230	.0188	.0153	.0124	.0100	.0081	.0065	.0052	.0042	.0033
1	.0979	.0847	.0730	.0626	.0535	.0455	.0385	.0325	.0273	.0228
2	.1952	.1792	.1635	.1482	.1336	.1198	.1068	.0947	.0835	.0732
3	.2421	.2359	.2279	.2185	.2079	.1964	.1843	.1718	.1591	.1465
4	.2092	.2162	.2212	.2242	.2252	.2243	.2215	.2171	.2112	.2040
5	.1334	.1464	.1586	.1699	.1802	.1891	.1964	.2026	.2071	.2099
6	.0650	.0757	.0869	.0984	.1101	.1218	.1333	.1445	.1551	.1649
7	.0247	.0305	.0371	.0444	.0524	.0611	.0704	.0803	.0905	.1010
8	.0074	.0097	.0125	.0158	.0197	.0242	.0293	.0351	.0416	.0487
9	.0017	.0024	.0033	.0044	.0058	.0075	.0096	.0121	.0151	.0185
10	.0003	.0005	.0007	.0010	.0014	.0019	.0025	.0033	.0043	.0056
11	.0000	.0001	.0001	.0002	.0002	.0004	.0005	.0007	.0010	.0013
12	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0001	.0002	.0002

r \ p	.31	.32	.33	.34	.35	.36	.37	.38	.39	.40
0	.0026	.0021	.0016	.0013	.0010	.0008	.0006	.0005	.0004	.0003
1	.0190	.0157	.0130	.0107	.0087	.0071	.0058	.0047	.0038	.0030
2	.0639	.0555	.0480	.0413	.0353	.0301	.0255	.0215	.0180	.0150
3	.1341	.1220	.1103	.0992	.0888	.0790	.0699	.0615	.0538	.0468
4	.1958	.1865	.1766	.1662	.1553	.1444	.1333	.1224	.1118	.1014
5	.2111	.2107	.2088	.2054	.2008	.1949	.1879	.1801	.1715	.1623
6	.1739	.1818	.1885	.1940	.1982	.2010	.2024	.2024	.2010	.1983
7	.1116	.1222	.1326	.1428	.1524	.1615	.1698	.1772	.1836	.1889
8	.0564	.0647	.0735	.0827	.0923	.1022	.1122	.1222	.1320	.1417
9	.0225	.0271	.0322	.0379	.0442	.0511	.0586	.0666	.0750	.0840
10	.0071	.0089	.0111	.0137	.0167	.0201	.0241	.0286	.0336	.0392
11	.0017	.0023	.0030	.0038	.0049	.0062	.0077	.0095	.0117	.0142
12	.0003	.0004	.0006	.0008	.0011	.0014	.0019	.0024	.0031	.0040
13	.0000	.0001	.0001	.0001	.0002	.0003	.0003	.0005	.0006	.0008
14	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0001

r \ p	.41	.42	.43	.44	.45	.46	.47	.48	.49	.50
0	.0002	.0002	.0001	.0001	.0001	.0001	.0000	.0000	.0000	.0000
1	.0024	.0019	.0015	.0012	.0009	.0007	.0005	.0004	.0003	.0002
2	.0125	.0103	.0085	.0069	.0056	.0046	.0037	.0029	.0023	.0018
3	.0405	.0349	.0299	.0254	.0215	.0181	.0151	.0126	.0104	.0085
4	.0915	.0821	.0732	.0649	.0572	.0501	.0436	.0378	.0325	.0278
5	.1526	.1426	.1325	.1224	.1123	.1024	.0929	.0837	.0749	.0667
6	.1944	.1894	.1833	.1762	.1684	.1600	.1510	.1416	.1319	.1222
7	.1930	.1959	.1975	.1978	.1969	.1947	.1912	.1867	.1811	.1746
8	.1509	.1596	.1676	.1742	.1812	.1865	.1908	.1939	.1958	.1964
9	.0932	.1027	.1124	.1221	.1318	.1413	.1504	.1591	.1672	.1746
10	.0453	.0521	.0594	.0672	.0755	.0842	.0934	.1028	.1124	.1222
11	.0172	.0206	.0244	.0288	.0337	.0391	.0452	.0518	.0589	.0667
12	.0050	.0062	.0077	.0094	.0115	.0139	.0167	.0199	.0236	.0278
13	.0011	.0014	.0018	.0023	.0029	.0036	.0046	.0057	.0070	.0085
14	.0002	.0002	.0003	.0004	.0005	.0007	.0009	.0011	.0014	.0018
15	.0000	.0000	.0000	.0000	.0001	.0001	.0001	.0001	.0002	.0002

تابع جدول (2)

n = 17

r \ p	.01	.02	.03	.04	.05	.06	.07	.08	.09	.10
0	.8429	.7093	.5958	.4996	.4181	.3493	.2912	.2423	.2012	.1668
1	.1447	.2461	.3133	.3539	.3741	.3790	.3726	.3582	.3383	.3150
2	.0117	.0402	.0775	.1180	.1575	.1935	.2244	.2492	.2677	.2800
3	.0006	.0041	.0120	.0246	.0415	.0618	.0844	.1083	.1324	.1556
4	.0000	.0003	.0013	.0036	.0076	.0138	.0222	.0330	.0458	.0605
5	.0000	.0000	.0001	.0004	.0010	.0023	.0044	.0075	.0118	.0175
6	.0000	.0000	.0000	.0000	.0001	.0003	.0007	.0013	.0023	.0039
7	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0002	.0004	.0007
8	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001

r \ p	.11	.12	.13	.14	.15	.16	.17	.18	.19	.20
0	.1379	.1138	.0937	.0770	.0631	.0516	.0421	.0343	.0278	.0225
1	.2898	.2638	.2381	.2131	.1893	.1671	.1466	.1279	.1109	.0957
2	.2865	.2878	.2846	.2775	.2673	.2547	.2402	.2245	.2081	.1914
3	.1771	.1963	.2126	.2259	.2359	.2425	.2460	.2464	.2441	.2393
4	.0766	.0937	.1112	.1287	.1457	.1617	.1784	.1893	.2004	.2093
5	.0246	.0332	.0432	.0545	.0668	.0801	.0939	.1081	.1222	.1361
6	.0061	.0091	.0129	.0177	.0236	.0305	.0385	.0474	.0573	.0680
7	.0012	.0019	.0030	.0045	.0065	.0091	.0124	.0164	.0211	.0267
8	.0002	.0003	.0006	.0009	.0014	.0022	.0032	.0045	.0062	.0084
9	.0000	.0000	.0001	.0002	.0003	.0004	.0006	.0010	.0015	.0021
10	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0002	.0003	.0004
11	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001

r \ p	.21	.22	.23	.24	.25	.26	.27	.28	.29	.30
0	.0182	.0146	.0118	.0094	.0075	.0060	.0047	.0038	.0030	.0023
1	.0822	.0702	.0597	.0505	.0426	.0357	.0299	.0248	.0206	.0169
2	.1747	.1584	.1427	.1277	.1136	.1005	.0883	.0772	.0672	.0581
3	.2322	.2234	.2131	.2016	.1893	.1765	.1634	.1502	.1372	.1245
4	.2161	.2205	.2228	.2228	.2209	.2170	.2115	.2044	.1961	.1868
5	.1493	.1617	.1730	.1830	.1914	.1982	.2033	.2067	.2083	.2081
6	.0794	.0912	.1034	.1156	.1278	.1393	.1504	.1608	.1701	.1784
7	.0332	.0404	.0485	.0573	.0668	.0769	.0874	.0982	.1093	.1201
8	.0110	.0143	.0181	.0226	.0279	.0338	.0404	.0478	.0558	.0644
9	.0029	.0040	.0054	.0071	.0093	.0119	.0150	.0186	.0228	.0276
10	.0006	.0009	.0013	.0018	.0025	.0033	.0044	.0058	.0074	.0095
11	.0001	.0002	.0002	.0004	.0005	.0007	.0010	.0014	.0019	.0026
12	.0000	.0000	.0000	.0001	.0001	.0001	.0002	.0003	.0004	.0006
13	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001

r \ p	.31	.32	.33	.34	.35	.36	.37	.38	.39	.40
0	.0018	.0014	.0011	.0009	.0007	.0005	.0004	.0003	.0002	.0002
1	.0139	.0114	.0093	.0075	.0060	.0048	.0039	.0031	.0024	.0019
2	.0500	.0428	.0364	.0309	.0160	.0211	.0182	.0151	.0125	.0102
3	.1123	.1007	.0898	.0795	.0701	.0614	.0534	.0463	.0394	.0341
4	.1766	.1659	.1547	.1434	.1320	.1208	.1099	.0993	.0892	.0796
5	.2063	.2030	.1982	.1921	.1849	.1761	.1677	.1582	.1482	.1379
6	.1854	.1910	.1952	.1979	.1991	.1888	.1970	.1939	.1895	.1839
7	.1309	.1413	.1511	.1602	.1685	.1757	.1818	.1868	.1904	.1927
8	.0735	.0831	.0930	.1032	.1134	.1235	.1335	.1431	.1521	.1606
9	.0330	.0391	.0458	.0531	.0611	.0695	.0784	.0877	.0973	.1070
10	.0119	.0147	.0181	.0218	.0263	.0313	.0368	.0430	.0498	.0571
11	.0034	.0044	.0057	.0072	.0090	.0112	.0138	.0168	.0202	.0242
12	.0008	.0010	.0014	.0018	.0024	.0031	.0040	.0051	.0065	.0081
13	.0001	.0002	.0003	.0004	.0005	.0007	.0009	.0012	.0016	.0021
14	.0000	.0000	.0000	.0001	.0001	.0001	.0002	.0002	.0003	.0004
15	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001

تابع جدول (2)

n = 17

r \ p	.41	.42	.43	.44	.45	.46	.47	.48	.49	.50
0	.0001	.0001	.0001	.0001	.0000	.0000	.0000	.0000	.0000	.0000
1	.0015	.0012	.0009	.0007	.0005	.0004	.0003	.0002	.0002	.0001
2	.0084	.0068	.0055	.0044	.0035	.0028	.0022	.0017	.0013	.0010
3	.0290	.0246	.0207	.0173	.0144	.0119	.0097	.0079	.0064	.0052
4	.0706	.0622	.0546	.0475	.0411	.0354	.0302	.0257	.0217	.0182
5	.1276	.1172	.1070	.0971	.0875	.0784	.0697	.0618	.0541	.0472
6	.1773	.1697	.1614	.1525	.1432	.1335	.1237	.1138	.1040	.0944
7	.1936	.1932	.1914	.1883	.1841	.1787	.1723	.1650	.1570	.1484
8	.1682	.1748	.1805	.1850	.1883	.1903	.1910	.1904	.1886	.1855
9	.1169	.1268	.1361	.1453	.1540	.1621	.1694	.1758	.1812	.1855
10	.0650	.0733	.0822	.0914	.1008	.1105	.1202	.1298	.1393	.1484
11	.0287	.0338	.0394	.0457	.0525	.0599	.0678	.0763	.0851	.0944
12	.0100	.0122	.0149	.0179	.0215	.0255	.0301	.0352	.0409	.0472
13	.0027	.0034	.0043	.0054	.0068	.0084	.0103	.0125	.0151	.0182
14	.0005	.0007	.0009	.0012	.0016	.0020	.0026	.0033	.0041	.0052
15	.0001	.0001	.0001	.0002	.0003	.0003	.0005	.0006	.0008	.0010
16	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0001	.0001

n = 18

r \ p	.01	.02	.03	.04	.05	.06	.07	.08	.09	.10
0	.8345	.6951	.5780	.4796	.3972	.3283	.2708	.2229	.1831	.1501
1	.1517	.2554	.3217	.3597	.3763	.3772	.3669	.3489	.3260	.3002
2	.0130	.0443	.0846	.1274	.1683	.2047	.2348	.2579	.2741	.2835
3	.0007	.0048	.0140	.0283	.0473	.0697	.0942	.1196	.1446	.1680
4	.0000	.0004	.0016	.0044	.0093	.0167	.0266	.0390	.0536	.0700
5	.0000	.0000	.0001	.0005	.0014	.0030	.0056	.0095	.0148	.0218
6	.0000	.0000	.0000	.0000	.0002	.0004	.0009	.0018	.0032	.0052
7	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0003	.0005	.0010
8	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0002

r \ p	.11	.12	.13	.14	.15	.16	.17	.18	.19	.20
0	.1227	.1002	.0815	.0662	.0536	.0434	.0349	.0281	.0225	.0180
1	.2731	.2458	.2193	.1940	.1704	.1486	.1288	.1110	.0951	.0811
2	.2869	.2850	.2785	.2685	.2556	.2407	.2243	.2071	.1897	.1723
3	.1891	.2072	.2220	.2331	.2406	.2445	.2450	.2425	.2373	.2297
4	.0877	.1060	.1244	.1423	.1592	.1746	.1882	.1996	.2087	.2153
5	.0303	.0405	.0520	.0648	.0787	.0931	.1079	.1227	.1371	.1507
6	.0081	.0120	.0168	.0229	.0301	.0384	.0479	.0584	.0697	.0816
7	.0017	.0028	.0043	.0064	.0091	.0126	.0168	.0220	.0280	.0350
8	.0003	.0005	.0009	.0014	.0022	.0033	.0047	.0066	.0090	.0120
9	.0000	.0001	.0001	.0003	.0004	.0007	.0011	.0016	.0024	.0033
10	.0000	.0000	.0000	.0000	.0001	.0001	.0002	.0003	.0005	.0008
11	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0001

r \ p	.21	.22	.23	.24	.25	.26	.27	.28	.29	.30
0	.0144	.0114	.0091	.0072	.0056	.0044	.0035	.0027	.0021	.0016
1	.0687	.0580	.0487	.0407	.0338	.0280	.0231	.0189	.0155	.0126
2	.1553	.1390	.1236	.1092	.0958	.0836	.0725	.0626	.0537	.0458
3	.2202	.1991	.1969	.1839	.1704	.1567	.1431	.1298	.1169	.1046
4	.2195	.2212	.2205	.2177	.2130	.2065	.1985	.1892	.1790	.1681
5	.1634	.1747	.1845	.1925	.1988	.2031	.2055	.2061	.2048	.2017
6	.0941	.1067	.1194	.1317	.1436	.1546	.1647	.1736	.1812	.1873
7	.0429	.0516	.0611	.0713	.0820	.0931	.1044	.1157	.1269	.1376
8	.0157	.0200	.0251	.0310	.0376	.0450	.0531	.0619	.0713	.0811
9	.0046	.0063	.0083	.0109	.0139	.0176	.0218	.0267	.0323	.0386
10	.0011	.0016	.0022	.0031	.0042	.0056	.0073	.0094	.0119	.0149
11	.0002	.0003	.0005	.0007	.0010	.0014	.0020	.0026	.0035	.0046
12	.0000	.0001	.0001	.0001	.0002	.0003	.0004	.0006	.0008	.0012
13	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0002	.0002

تابع جدول (2)

n = 18

r \ p	.31	.32	.33	.34	.35	.36	.37	.38	.39	.40
0	.0013	.0010	.0007	.0006	.0004	.0003	.0002	.0002	.0001	.0001
1	.0102	.0082	.0066	.0052	.0042	.0033	.0026	.0020	.0016	.0012
2	.0388	.0327	.0275	.0229	.0190	.0157	.0120	.0105	.0085	.0069
3	.0930	.0822	.0722	.0630	.0547	.0471	.0404	.0344	.0292	.0246
4	.1567	.1450	.1333	.1217	.1104	.0994	.0850	.0791	.0699	.0614
5	.1971	.1911	.1838	.1755	.1664	.1566	.1463	.1358	.1252	.1146
6	.1919	.1948	.1962	.1959	.1941	.1908	.1862	.1803	.1734	.1655
7	.1478	.1572	.1656	.1730	.1792	.1840	.1875	.1895	.1900	.1892
8	.0913	.1017	.1122	.1226	.1327	.1423	.1514	.1597	.1671	.1734
9	.0456	.0532	.0614	.0701	.0794	.0890	.0988	.1087	.1187	.1284
10	.0184	.0225	.0272	.0325	.0385	.0450	.0522	.0600	.0683	.0771
11	.0060	.0077	.0097	.0122	.0151	.0184	.0223	.0267	.0318	.0374
12	.0016	.0021	.0028	.0037	.0047	.0060	.0076	.0096	.0118	.0145
13	.0003	.0005	.0006	.0009	.0012	.0016	.0021	.0027	.0035	.0045
14	.0001	.0001	.0001	.0002	.0002	.0003	.0004	.0006	.0008	.0011
15	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0001	.0002

r \ p	.41	.42	.43	.44	.45	.46	.47	.48	.49	.50
0	.0001	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
1	.0009	.0007	.0005	.0004	.0003	.0002	.0002	.0001	.0001	.0001
2	.0055	.0044	.0035	.0028	.0022	.0017	.0013	.0010	.0008	.0006
3	.0206	.0171	.0141	.0116	.0095	.0077	.0062	.0050	.0039	.0031
4	.0536	.0464	.0400	.0342	.0291	.0246	.0206	.0172	.0142	.0117
5	.1042	.0941	.0844	.0753	.0666	.0586	.0512	.0444	.0382	.0327
6	.1569	.1477	.1380	.1281	.1181	.1081	.0983	.0887	.0796	.0708
7	.1869	.1833	.1785	.1726	.1657	.1579	.1494	.1404	.1310	.1214
8	.1786	.1825	.1852	.1864	.1864	.1850	.1822	.1782	.1731	.1669
9	.1379	.1469	.1552	.1628	.1694	.1751	.1795	.1828	.1848	.1855
10	.0862	.0957	.1054	.1151	.1248	.1342	.1433	.1519	.1598	.1669
11	.0436	.0504	.0578	.0658	.0742	.0831	.0924	.1020	.1117	.1214
12	.0177	.0213	.0254	.0301	.0354	.0413	.0478	.0549	.0626	.0708
13	.0057	.0071	.0089	.0109	.0134	.0162	.0196	.0234	.0278	.0327
14	.0014	.0018	.0024	.0031	.0039	.0049	.0062	.0077	.0095	.0117
15	.0003	.0004	.0005	.0006	.0009	.0011	.0015	.0019	.0024	.0031
16	.0000	.0000	.0001	.0001	.0001	.0002	.0002	.0003	.0004	.0006
17	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001

n = 19

r \ p	.01	.02	.03	.04	.05	.06	.07	.08	.09	.10
0	.8262	.6812	.5606	.4604	.3774	.3086	.2519	.2051	.1666	.1351
1	.1586	.2642	.3294	.3645	.3774	.3743	.3602	.3389	.3131	.2852
2	.0144	.0485	.0917	.1367	.1787	.2150	.2440	.2652	.2787	.2852
3	.0008	.0056	.0161	.0323	.0533	.0778	.1041	.1307	.1562	.1796
4	.0000	.0005	.0020	.0054	.0112	.0199	.0313	.0455	.0618	.0798
5	.0000	.0000	.0002	.0007	.0018	.0038	.0071	.0119	.0183	.0266
6	.0000	.0000	.0000	.0001	.0002	.0006	.0012	.0024	.0042	.0069
7	.0000	.0000	.0000	.0000	.0000	.0001	.0002	.0004	.0008	.0014
8	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0002

r \ p	.11	.12	.13	.14	.15	.16	.17	.18	.19	.20
0	.1092	.0881	.0709	.0569	.0456	.0364	.0290	.0230	.0182	.0144
1	.2565	.2284	.2014	.1761	.1529	.1318	.1129	.0961	.0813	.0685
2	.2854	.2803	.2708	.2581	.2428	.2259	.2081	.1898	.1717	.1540
3	.1999	.2166	.2293	.2381	.2428	.2439	.2415	.2361	.2282	.2182
4	.0988	.1181	.1371	.1550	.1714	.1858	.1979	.2073	.2141	.2182
5	.0366	.0483	.0614	.0757	.0907	.1062	.1216	.1365	.1507	.1636
6	.0106	.0154	.0214	.0288	.0374	.0472	.0581	.0699	.0825	.0955
7	.0024	.0039	.0059	.0087	.0122	.0167	.0221	.0285	.0359	.0443
8	.0004	.0008	.0013	.0021	.0032	.0048	.0068	.0094	.0126	.0166
9	.0001	.0001	.0002	.0004	.0007	.0011	.0017	.0025	.0036	.0051
10	.0000	.0000	.0000	.0001	.0001	.0002	.0003	.0006	.0009	.0013
11	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0002	.0003

تابع جدول (2)

n = 19

r \ p	.21	.22	.23	.24	.25	.26	.27	.28	.29	.30
0	.0113	.0087	.0070	.0054	.0042	.0033	.0025	.0019	.0015	.0011
1	.0573	.0477	.0396	.0326	.0268	.0219	.0178	.0144	.0116	.0093
2	.1371	.1212	.1064	.0927	.0803	.0692	.0592	.0503	.0426	.0358
3	.2065	.1937	.1800	.1659	.1517	.1377	.1240	.1109	.0985	.0869
4	.2196	.2185	.2151	.2096	.2023	.1935	.1835	.1726	.1610	.1491
5	.1751	.1849	.1928	.1986	.2023	.2040	.2036	.2013	.1973	.1916
6	.1086	.1217	.1343	.1463	.1574	.1672	.1757	.1827	.1880	.1916
7	.0536	.0637	.0745	.0858	.0974	.1091	.1207	.1320	.1426	.1525
8	.0211	.0270	.0334	.0406	.0487	.0575	.0670	.0770	.0874	.0981
9	.0069	.0093	.0122	.0157	.0198	.0247	.0303	.0366	.0436	.0514
10	.0018	.0026	.0036	.0050	.0066	.0087	.0112	.0142	.0178	.0220
11	.0004	.0006	.0009	.0013	.0018	.0025	.0034	.0045	.0060	.0077
12	.0001	.0001	.0002	.0003	.0004	.0006	.0008	.0012	.0016	.0022
13	.0000	.0000	.0000	.0000	.0001	.0001	.0002	.0002	.0004	.0005
14	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001

r \ p	.31	.32	.33	.34	.35	.36	.37	.38	.39	.40
0	.0009	.0007	.0005	.0004	.0003	.0002	.0002	.0001	.0001	.0001
1	.0074	.0059	.0046	.0036	.0029	.0022	.0017	.0013	.0010	.0008
2	.0299	.0249	.0206	.0169	.0138	.0112	.0091	.0073	.0058	.0046
3	.0762	.0664	.0574	.0494	.0422	.0358	.0302	.0253	.0211	.0175
4	.1370	.1247	.1131	.1017	.0909	.0806	.0710	.0621	.0540	.0467
5	.1846	.1764	.1672	.1572	.1468	.1360	.1251	.1143	.1036	.0933
6	.1935	.1936	.1921	.1890	.1844	.1785	.1714	.1634	.1546	.1451
7	.1615	.1692	.1757	.1808	.1844	.1865	.1870	.1860	.1835	.1797
8	.1088	.1195	.1298	.1397	.1489	.1573	.1647	.1710	.1760	.1797
9	.0597	.0687	.0762	.0880	.0980	.1082	.1182	.1281	.1375	.1464
10	.0268	.0323	.0385	.0453	.0528	.0608	.0694	.0785	.0879	.0976
11	.0099	.0124	.0155	.0191	.0233	.0280	.0334	.0394	.0460	.0532
12	.0030	.0039	.0051	.0066	.0083	.0105	.0131	.0161	.0196	.0237
13	.0007	.0010	.0014	.0018	.0024	.0032	.0041	.0053	.0067	.0085
14	.0001	.0002	.0003	.0004	.0006	.0008	.0010	.0014	.0018	.0024
15	.0000	.0000	.0000	.0001	.0001	.0001	.0002	.0003	.0004	.0005
16	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001

r \ p	.41	.42	.43	.44	.45	.46	.47	.48	.49	.50
0	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
1	.0006	.0004	.0003	.0002	.0002	.0001	.0001	.0001	.0001	.0000
2	.0037	.0025	.0022	.0017	.0013	.0010	.0008	.0006	.0004	.0003
3	.0144	.0118	.0096	.0077	.0062	.0049	.0039	.0031	.0024	.0018
4	.0400	.0341	.0289	.0243	.0203	.0168	.0138	.0113	.0092	.0074
5	.0834	.0741	.0653	.0572	.0497	.0429	.0368	.0313	.0265	.0222
6	.1353	.1252	.1150	.1049	.0949	.0853	.0751	.0674	.0593	.0518
7	.1746	.1683	.1611	.1530	.1443	.1350	.1254	.1156	.1058	.0961
8	.1820	.1829	.1823	.1803	.1771	.1725	.1668	.1601	.1525	.1442
9	.1546	.1618	.1681	.1732	.1771	.1796	.1808	.1806	.1791	.1762
10	.1074	.1172	.1268	.1361	.1449	.1530	.1603	.1667	.1721	.1762
11	.0611	.0694	.0783	.0875	.0970	.1066	.1163	.1259	.1352	.1442
12	.0283	.0335	.0394	.0458	.0529	.0606	.0688	.0775	.0866	.0961
13	.0106	.0131	.0160	.0194	.0233	.0278	.0328	.0385	.0448	.0518
14	.0032	.0041	.0052	.0065	.0082	.0101	.0125	.0152	.0185	.0222
15	.0007	.0010	.0013	.0017	.0022	.0029	.0037	.0047	.0059	.0074
16	.0001	.0002	.0002	.0003	.0005	.0006	.0008	.0011	.0014	.0018
17	.0000	.0000	.0000	.0000	.0001	.0001	.0001	.0002	.0002	.0003

n = 20

r \ p	.01	.02	.03	.04	.05	.06	.07	.08	.09	.10
0	.8179	.6676	.5438	.4420	.3585	.2901	.2342	.1887	.1516	.1216
1	.1652	.2725	.3364	.3683	.3774	.3703	.3526	.3282	.3000	.2702
2	.0159	.0528	.0988	.1458	.1887	.2246	.2521	.2711	.2818	.2852
3	.0010	.0065	.0183	.0364	.0596	.0860	.1139	.1414	.1672	.1901
4	.0000	.0006	.0024	.0065	.0133	.0233	.0364	.0523	.0703	.0898
5	.0000	.0000	.0002	.0009	.0022	.0048	.0088	.0145	.0222	.0319
6	.0000	.0000	.0000	.0001	.0003	.0008	.0017	.0032	.0055	.0089
7	.0000	.0000	.0000	.0000	.0000	.0001	.0002	.0005	.0011	.0020
8	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0002	.0004
9	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001

تابع جدول (2)

n = 20

r \ p	.11	.12	.13	.14	.15	.16	.17	.18	.19	.20
0	.0972	.0776	.0617	.0490	.0388	.0306	.0241	.0189	.0148	.0115
1	.2403	.2115	.1844	.1595	.1368	.1165	.0986	.0829	.0693	.0576
2	.2822	.2740	.2618	.2466	.2293	.2109	.1919	.1730	.1545	.1369
3	.2093	.2242	.2347	.2409	.2428	.2410	.2358	.2278	.2175	.2054
4	.1099	.1299	.1491	.1666	.1821	.1951	.2053	.2125	.2168	.2182
5	.0435	.0567	.0713	.0868	.1028	.1189	.1345	.1493	.1627	.1746
6	.0134	.0193	.0268	.0353	.0454	.0566	.0689	.0819	.0954	.1091
7	.0033	.0053	.0080	.0115	.0160	.0216	.0282	.0360	.0448	.0545
8	.0007	.0012	.0019	.0030	.0046	.0067	.0094	.0128	.0171	.0222
9	.0001	.0002	.0004	.0007	.0011	.0017	.0026	.0038	.0053	.0074
10	.0000	.0000	.0001	.0001	.0002	.0004	.0006	.0009	.0014	.0020
11	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0002	.0003	.0005
12	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001

r \ p	.21	.22	.23	.24	.25	.26	.27	.28	.29	.30
0	.0090	.0069	.0054	.0041	.0032	.0024	.0018	.0014	.0011	.0008
1	.0477	.0392	.0321	.0261	.0211	.0170	.0137	.0109	.0087	.0068
2	.1204	.1050	.0910	.0783	.0669	.0569	.0480	.0403	.0336	.0278
3	.1920	.1777	.1631	.1484	.1339	.1199	.1065	.0940	.0823	.0716
4	.2169	.2131	.2070	.1991	.1897	.1790	.1675	.1553	.1429	.1304
5	.1845	.1923	.1979	.2012	.2023	.2013	.1982	.1933	.1868	.1789
6	.1226	.1356	.1478	.1589	.1686	.1768	.1833	.1879	.1907	.1916
7	.0652	.0765	.0883	.1003	.1124	.1242	.1356	.1462	.1558	.1643
8	.0282	.0351	.0429	.0515	.0609	.0709	.0815	.0924	.1034	.1144
9	.0100	.0132	.0171	.0217	.0271	.0332	.0402	.0479	.0563	.0654
10	.0029	.0041	.0056	.0075	.0099	.0128	.0163	.0205	.0253	.0308
11	.0007	.0010	.0015	.0022	.0030	.0041	.0055	.0072	.0094	.0120
12	.0001	.0002	.0003	.0005	.0008	.0011	.0015	.0021	.0029	.0039
13	.0000	.0000	.0001	.0001	.0002	.0002	.0003	.0005	.0007	.0010
14	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0001	.0002

r \ p	.31	.32	.33	.34	.35	.36	.37	.38	.39	.40
0	.0006	.0004	.0003	.0002	.0002	.0001	.0001	.0001	.0001	.0000
1	.0054	.0042	.0033	.0025	.0020	.0015	.0011	.0009	.0007	.0005
2	.0229	.0188	.0153	.0124	.0100	.0080	.0064	.0050	.0040	.0031
3	.0619	.0531	.0453	.0383	.0323	.0270	.0224	.0185	.0152	.0123
4	.1181	.1062	.0947	.0839	.0738	.0645	.0559	.0482	.0412	.0350
5	.1698	.1599	.1493	.1384	.1272	.1161	.1051	.0945	.0843	.0746
6	.1907	.1881	.1839	.1782	.1712	.1632	.1543	.1447	.1347	.1244
7	.1714	.1770	.1811	.1836	.1844	.1836	.1812	.1774	.1722	.1659
8	.1251	.1354	.1450	.1537	.1614	.1678	.1730	.1767	.1790	.1797
9	.0750	.0849	.0952	.1056	.1158	.1259	.1354	.1444	.1526	.1597
10	.0370	.0440	.0516	.0598	.0686	.0779	.0875	.0974	.1073	.1171
11	.0151	.0188	.0231	.0280	.0336	.0398	.0467	.0542	.0624	.0710
12	.0051	.0066	.0085	.0108	.0136	.0168	.0206	.0249	.0299	.0355
13	.0014	.0019	.0026	.0034	.0045	.0058	.0074	.0094	.0118	.0146
14	.0003	.0005	.0006	.0009	.0012	.0016	.0022	.0029	.0038	.0049
15	.0001	.0001	.0001	.0002	.0003	.0004	.0005	.0007	.0010	.0013
16	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0001	.0002	.0003

r \ p	.41	.42	.43	.44	.45	.46	.47	.48	.49	.50
0	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
1	.0004	.0003	.0002	.0001	.0001	.0001	.0001	.0000	.0000	.0000
2	.0024	.0018	.0014	.0011	.0008	.0006	.0005	.0003	.0002	.0002
3	.0100	.0080	.0064	.0051	.0040	.0031	.0024	.0018	.0014	.0011
4	.0295	.0247	.0206	.0170	.0139	.0113	.0092	.0074	.0059	.0046
5	.0656	.0573	.0496	.0423	.0365	.0309	.0260	.0217	.0180	.0148
6	.1140	.1037	.0936	.0839	.0746	.0658	.0577	.0501	.0432	.0370
7	.1585	.1502	.1413	.1318	.1221	.1122	.1023	.0925	.0830	.0739
8	.1790	.1768	.1732	.1683	.1623	.1553	.1474	.1388	.1296	.1201
9	.1658	.1707	.1742	.1763	.1771	.1763	.1742	.1708	.1661	.1602
10	.1268	.1359	.1446	.1524	.1593	.1652	.1700	.1734	.1755	.1762
11	.0801	.0895	.0991	.1089	.1185	.1280	.1370	.1455	.1533	.1602
12	.0417	.0486	.0561	.0642	.0727	.0818	.0911	.1007	.1105	.1201
13	.0178	.0217	.0260	.0310	.0366	.0429	.0497	.0572	.0653	.0739
14	.0062	.0078	.0098	.0122	.0150	.0183	.0221	.0264	.0314	.0370
15	.0017	.0023	.0030	.0038	.0049	.0062	.0078	.0098	.0121	.0148
16	.0004	.0005	.0007	.0009	.0013	.0017	.0022	.0028	.0036	.0046
17	.0001	.0001	.0001	.0002	.0003	.0003	.0005	.0006	.0008	.0011
18	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0001	.0002

تابع جدول (2)

n = 25

r \ p	.01	.02	.03	.04	.05	.06	.07	.08	.09	.10
0	7778	6035	4670	3604	2774	2129	1630	1244	0946	0718
1	1964	3079	3611	3754	3650	3398	3066	2704	2340	1994
2	0238	0754	1340	1877	2305	2602	2770	2821	2777	2659
3	0018	0118	0318	0600	0930	1273	1598	1881	2106	2265
4	0001	0013	0054	0137	0269	0447	0662	0899	1145	1384
5	0000	0001	0007	0024	0060	0120	0209	0329	0476	0646
6	0000	0000	0001	0003	0010	0026	0052	0095	0157	0239
7	0000	0000	0000	0000	0001	0004	0011	0022	0042	0072
8	0000	0000	0000	0000	0000	0001	0002	0004	0009	0018
9	0	0000	0000	0000	0000	0000	0000	0001	0002	0004
10	0	0000	0000	0000	0000	0000	0000	0000	0000	0001

r \ p	.11	.12	.13	.14	.15	.16	.17	.18	.19	.20
0	0543	0409	0308	0230	0172	0128	0095	0070	0052	0038
1	1678	1395	1149	0938	0759	0609	0486	0384	0302	0236
2	2488	2283	2060	1832	1607	1392	1193	1012	0851	0708
3	2358	2387	2360	2286	2174	2033	1874	1704	1530	1358
4	1603	1790	1940	2047	2110	2130	2111	2057	1974	1867
5	0832	1025	1217	1399	1564	1704	1816	1897	1945	1960
6	0343	0466	0606	0759	0920	1082	1240	1388	1520	1633
7	0115	0173	0246	0336	0441	0559	0689	0827	0968	1108
8	0032	0053	0083	0123	0175	0240	0318	0408	0511	0623
9	0007	0014	0023	0038	0058	0086	0123	0169	0226	0294
10	0001	0003	0006	0010	0016	0026	0040	0059	0085	0118
11	0000	0001	0001	0002	0004	0007	0011	0018	0027	0040
12	0000	0000	0000	0000	0001	0002	0003	0005	0007	0012
13	0000	0000	0000	0000	0000	0000	0001	0001	0002	0003
14	0000	0000	0000	0000	0000	0000	0000	0000	0000	0001

r \ p	.21	.22	.23	.24	.25	.26	.27	.28	.29	.30
0	0028	0020	0015	0010	0008	0005	0004	0003	0002	0001
1	0183	0141	0109	0083	0063	0047	0035	0026	0020	0014
2	0585	0479	0389	0314	0251	0199	0157	0123	0096	0074
3	1192	1035	0891	0759	0641	0537	0446	0367	0300	0243
4	1742	1606	1463	1318	1175	1037	0906	0785	0673	0572
5	1945	1903	1836	1749	1645	1531	1408	1282	1155	1030
6	1724	1789	1828	1841	1828	1793	1736	1661	1572	1472
7	1244	1369	1482	1578	1654	1709	1743	1754	1743	1712
8	0744	0869	0996	1121	1241	1351	1450	1535	1602	1651
9	0373	0463	0562	0669	0781	0897	1013	1127	1236	1336
10	0159	0209	0269	0338	0417	0504	0600	0701	0808	0916
11	0058	0080	0109	0145	0189	0242	0302	0372	0450	0536
12	0018	0026	0038	0054	0074	0099	0130	0169	0214	0268
13	0005	0007	0011	0017	0025	0035	0048	0066	0088	0115
14	0001	0002	0003	0005	0007	0010	0015	0022	0031	0042
15	0000	0000	0001	0001	0002	0003	0004	0006	0009	0013
16	0000	0000	0000	0000	0000	0001	0001	0002	0002	0004
17	0000	0000	0000	0000	0000	0000	0000	0000	0001	0001

r \ p	.31	.32	.33	.34	.35	.36	.37	.38	.39	.40
0	0001	0001	0000	0000	0000	0000	0000	0000	0000	0000
1	0011	0008	0006	0004	0003	0002	0001	0001	0001	0000
2	0057	0043	0033	0025	0018	0014	0010	0007	0005	0004
3	0195	0156	0123	0097	0076	0058	0045	0034	0026	0019
4	0482	0403	0334	0274	0224	0181	0145	0115	0091	0071
5	0910	0797	0691	0594	0506	0427	0357	0297	0244	0199
6	1363	1250	1134	1020	0908	0801	0700	0606	0520	0442
7	1662	1596	1516	1426	1327	1222	1115	1008	0902	0800
8	1680	1690	1681	1652	1607	1547	1474	1390	1298	1200
9	1426	1502	1563	1608	1635	1644	1635	1609	1567	1511
10	1025	1131	1232	1325	1409	1479	1536	1578	1603	1612
11	0628	0726	0828	0931	1034	1135	1230	1319	1398	1465
12	0329	0399	0476	0560	0650	0745	0843	0943	1043	1140
13	0148	0188	0234	0288	0350	0419	0495	0578	0667	0760
14	0057	0076	0099	0127	0161	0202	0249	0304	0365	0434
15	0019	0026	0036	0048	0064	0083	0107	0136	0171	0212
16	0005	0008	0011	0015	0021	0029	0039	0052	0068	0088
17	0001	0002	0003	0004	0006	0009	0012	0017	0023	0031
18	0000	0000	0001	0001	0001	0002	0003	0005	0007	0009
19	0000	0000	0000	0000	0000	0000	0001	0001	0002	0002

تابع جدول (2)

r \ p	.41	.42	.43	.44	.45	.46	.47	.48	.49	.50
	n = 25									
0	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000
1	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000
2	0003	0002	0001	0001	0001	0000	0000	0000	0000	0000
3	0014	0011	0008	0006	0004	0003	0002	0001	0001	0001
4	0055	0042	0033	0024	0018	0014	0010	0007	0005	0004
5	0161	0129	0102	0081	0063	0049	0037	0028	0021	0016
6	0372	0311	0257	0211	0172	0138	0110	0087	0068	0053
7	0703	0611	0527	0450	0381	0319	0265	0218	0178	0143
8	1099	0996	0895	0796	0701	0612	0521	0453	0369	0322
9	1442	1363	1275	1181	1084	0985	0886	0790	0637	0609
10	1603	1576	1538	1485	1419	1342	1257	1169	1011	0974
11	1519	1559	1583	1591	1583	1559	1521	1468	1404	1328
12	1232	1317	1393	1458	1511	1550	1573	1581	1573	1550
13	0856	0954	1051	1146	1236	1320	1395	1460	1512	1550
14	0510	0592	0680	0772	0867	0964	1060	1155	1245	1328
15	0260	0314	0376	0445	0520	0602	0690	0782	0871	0974
16	0113	0142	0177	0218	0266	0321	0382	0451	0527	0609
17	0042	0055	0071	0091	0115	0145	0179	0220	0258	0322
18	0013	0018	0024	0032	0042	0055	0071	0090	0114	0143
19	0003	0005	0007	0009	0013	0017	0023	0031	0040	0053
20	0001	0001	0001	0002	0003	0004	0006	0009	0012	0016
21	0000	0000	0000	0000	0001	0001	0001	0002	0003	0004
22	0000	0000	0000	0000	0000	0000	0000	0000	0000	0001

جدول (3) : التوزيع الاحتمالي لتوزيع بواسون لقيم مختلفة بالنسبة الى λ ، حيث

$$p_X(x) = P(X = x) = \frac{e^{-\lambda}\,\lambda^{x}}{x!} \quad , x = 0, 1, 2, 3, \cdots$$

x	λ = .1	.2	.3	.4	.5	.6	.7	.8	.9	1.0
0	.9048	.8187	.7408	.6703	.6065	.5488	.4966	.4493	.4066	.3679
1	.0905	.1637	.2222	.2681	.3033	.3293	.3476	.3595	.3659	.3679
2	.0045	.0164	.0333	.0536	.0758	.0988	.1217	.1438	.1647	.1839
3	.0002	.0011	.0033	.0072	.0126	.0198	.0284	.0383	.0494	.0613
4	.0000	.0001	.0002	.0007	.0016	.0030	.0050	.0077	.0111	.0153
5	.0000	.0000	.0000	.0001	.0002	.0004	.0007	.0012	.0020	.0031
6	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0002	.0003	.0005
7	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001

x	λ = 1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2.0
0	.3329	.3012	.2725	.2466	.2231	.2019	.1827	.1653	.1496	.1353
1	.3662	.3614	.3543	.3452	.3347	.3230	.3106	.2975	.2842	.2707
2	.2014	.2169	.2303	.2417	.2510	.2584	.2640	.2678	.2700	.2707
3	.0738	.0867	.0998	.1128	.1255	.1378	.1496	.1607	.1710	.1804
4	.0203	.0260	.0324	.0395	.0471	.0551	.0636	.0723	.0812	.0902
5	.0045	.0062	.0084	.0111	.0141	.0176	.0216	.0260	.0309	.0361
6	.0008	.0012	.0018	.0026	.0035	.0047	.0061	.0078	.0098	.0120
7	.0001	.0002	.0003	.0005	.0008	.0011	.0015	.0020	.0027	.0034
8	.0000	.0000	.0001	.0001	.0001	.0002	.0003	.0005	.0006	.0009
9	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0001	.0002

x	λ = 2.1	2.2	2.3	2.4	2.5	2.6	2.7	2.8	2.9	3.0
0	.1225	.1108	.1003	.0907	.0821	.0743	.0672	.0608	.0550	.0498
1	.2572	.2438	.2306	.2177	.2052	.1931	.1815	.1703	.1596	.1494
2	.2700	.2681	.2652	.2613	.2565	.2510	.2450	.2384	.2314	.2240
3	.1890	.1966	.2033	.2090	.2138	.2176	.2205	.2225	.2237	.2240
4	.0992	.1082	.1169	.1254	.1336	.1414	.1488	.1557	.1622	.1680
5	.0417	.0476	.0538	.0602	.0668	.0735	.0804	.0872	.0940	.1008
6	.0146	.0174	.0206	.0241	.0278	.0319	.0362	.0407	.0455	.0504
7	.0044	.0055	.0068	.0083	.0099	.0118	.0139	.0163	.0188	.0216
8	.0011	.0015	.0019	.0025	.0031	.0038	.0047	.0057	.0068	.0081
9	.0003	.0004	.0005	.0007	.0009	.0011	.0014	.0018	.0022	.0027
10	.0001	.0001	.0001	.0002	.0002	.0003	.0004	.0005	.0006	.0008
11	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0001	.0002	.0002
12	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001

x	λ = 3.1	3.2	3.3	3.4	3.5	3.6	3.7	3.8	3.9	4.0
0	.0450	.0408	.0369	.0334	.0302	.0273	.0247	.0224	.0202	.0183
1	.1397	.1304	.1217	.1135	.1057	.0984	.0915	.0850	.0789	.0733
2	.2165	.2087	.2008	.1929	.1850	.1771	.1692	.1615	.1539	.1465
3	.2237	.2226	.2209	.2186	.2158	.2125	.2087	.2046	.2001	.1954
4	.1734	.1781	.1823	.1858	.1888	.1912	.1931	.1944	.1951	.1954
5	.1075	.1140	.1203	.1264	.1322	.1377	.1429	.1477	.1522	.1563
6	.0555	.0608	.0662	.0716	.0771	.0826	.0881	.0936	.0989	.1042
7	.0246	.0278	.0312	.0348	.0385	.0425	.0466	.0508	.0551	.0595
8	.0095	.0111	.0129	.0148	.0169	.0191	.0215	.0241	.0269	.0298
9	.0033	.0040	.0047	.0056	.0066	.0076	.0089	.0102	.0116	.0132
10	.0010	.0013	.0016	.0019	.0023	.0028	.0033	.0039	.0045	.0053
11	.0003	.0004	.0005	.0006	.0007	.0009	.0011	.0013	.0016	.0019
12	.0001	.0001	.0001	.0002	.0002	.0003	.0003	.0004	.0005	.0006
13	.0000	.0000	.0000	.0000	.0001	.0001	.0001	.0001	.0002	.0002
14	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001

المصدر :

Handbook of Probability and Statistics With Tables 2nd ed. McGraw Hill Inc.1970.

تابع جدول (3)

x	λ = 4.1	4.2	4.3	4.4	4.5	4.6	4.7	4.8	4.9	5.0
0	0166	0150	0136	0123	0111	0101	0091	0082	0074	0067
1	0679	0630	0583	0540	0500	0462	0427	0395	0365	0337
2	1393	1323	1254	1188	1125	1063	1005	0948	0894	0842
3	1904	1852	1798	1743	1687	1631	1574	1517	1460	1404
4	1951	1944	1933	1917	1898	1875	1849	1820	1789	1755
5	1600	1633	1662	1687	1708	1725	1738	1747	1753	1755
6	1093	1143	1191	1237	1281	1323	1362	1398	1432	1462
7	0640	0686	0732	0778	0824	0869	0914	0959	1002	1044
8	0328	0360	0393	0428	0463	0500	0537	0575	0614	0653
9	0150	0168	0188	0209	0232	0255	0280	0307	0334	0363
10	0061	0071	0081	0092	0104	0118	0132	0147	0164	0181
11	0023	0027	0032	0037	0043	0049	0056	0064	0073	0082
12	0008	0009	0011	0014	0016	0019	0022	0026	0030	0034
13	0002	0003	0004	0005	0006	0007	0008	0009	0011	0013
14	0001	0001	0001	0001	0002	0002	0003	0003	0004	0005
15	0000	0000	0000	0000	0001	0001	0001	0001	0001	0002

x	λ = 5.1	5.2	5.3	5.4	5.5	5.6	5.7	5.8	5.9	6.0
0	0061	0055	0050	0045	0041	0037	0033	0030	0027	0025
1	0311	0287	0265	0244	0225	0207	0191	0176	0162	0149
2	0793	0746	0701	0659	0618	0580	0544	0509	0477	0446
3	1348	1293	1239	1185	1133	1082	1033	0985	0938	0892
4	1719	1681	1641	1600	1558	1515	1472	1428	1383	1339
5	1753	1748	1740	1728	1714	1697	1678	1656	1632	1606
6	1490	1515	1537	1555	1571	1584	1594	1601	1605	1606
7	1086	1125	1163	1200	1234	1267	1298	1326	1353	1377
8	0692	0731	0771	0810	0849	0887	0925	0962	0998	1033
9	0392	0423	0454	0486	0519	0552	0586	0620	0654	0688
10	0200	0220	0241	0262	0285	0309	0334	0359	0386	0413
11	0093	0104	0116	0129	0143	0157	0173	0190	0207	0225
12	0039	0045	0051	0058	0065	0073	0082	0092	0102	0113
13	0015	0018	0021	0024	0028	0032	0036	0041	0046	0052
14	0006	0007	0008	0009	0011	0013	0015	0017	0019	0022
15	0002	0002	0003	0003	0004	0005	0006	0007	0008	0009
16	0001	0001	0001	0001	0001	0002	0002	0002	0003	0003
17	0000	0000	0000	0000	0000	0001	0001	0001	0001	0001

x	λ = 6.1	6.2	6.3	6.4	6.5	6.6	6.7	6.8	6.9	7.0
0	0022	0020	0018	0017	0015	0014	0012	0011	0010	0009
1	0137	0126	0116	0106	0098	0090	0082	0076	0070	0064
2	0417	0390	0364	0340	0318	0296	0276	0258	0240	0223
3	0848	0806	0765	0726	0688	0652	0617	0584	0552	0521
4	1294	1249	1205	1162	1118	1076	1034	0992	0952	0912
5	1579	1549	1519	1487	1454	1420	1385	1349	1314	1277
6	1605	1601	1595	1586	1575	1562	1546	1529	1511	1490
7	1399	1418	1435	1450	1462	1472	1480	1486	1489	1490
8	1066	1099	1130	1160	1188	1215	1240	1263	1284	1304
9	0723	0757	0791	0825	0858	0891	0923	0954	0985	1014
10	0441	0469	0498	0528	0558	0588	0618	0649	0679	0710
11	0245	0265	0285	0307	0330	0353	0377	0401	0426	0452
12	0124	0137	0150	0164	0179	0194	0210	0227	0245	0264
13	0058	0065	0073	0081	0089	0099	0108	0119	0130	0142
14	0025	0029	0033	0037	0041	0046	0052	0058	0064	0071
15	0010	0012	0014	0016	0018	0020	0023	0026	0029	0033
16	0004	0005	0005	0006	0007	0008	0010	0011	0013	0014
17	0001	0002	0002	0002	0003	0003	0004	0004	0005	0006
18	0000	0001	0001	0001	0001	0001	0001	0002	0002	0002
19	0000	0000	0000	0000	0000	0000	0000	0001	0001	0001

تابع جدول (3)

x	λ 7.1	7.2	7.3	7.4	7.5	7.6	7.7	7.8	7.9	8.0
0	.0008	.0007	.0007	.0006	.0006	.0005	.0005	.0004	.0004	.0003
1	.0059	.0054	.0049	.0045	.0041	.0038	.0035	.0032	.0029	.0027
2	.0208	.0194	.0180	.0167	.0156	.0145	.0134	.0125	.0116	.0107
3	.0492	.0464	.0438	.0413	.0389	.0366	.0345	.0324	.0305	.0286
4	.0874	.0836	.0799	.0764	.0729	.0696	.0663	.0632	.0602	.0573
5	.1241	.1204	.1167	.1130	.1094	.1057	.1021	.0986	.0951	.0916
6	.1468	.1445	.1420	.1394	.1367	.1339	.1311	.1282	.1252	.1221
7	.1489	.1486	.1481	.1474	.1465	.1454	.1442	.1428	.1413	.1396
8	.1321	.1337	.1351	.1363	.1373	.1382	.1388	.1392	.1395	.1396
9	.1042	.1070	.1096	.1121	.1144	.1167	.1187	.1207	.1224	.1241
10	.0740	.0770	.0800	.0829	.0858	.0887	.0914	.0941	.0967	.0993
11	.0478	.0504	.0531	.0558	.0585	.0613	.0640	.0667	.0695	.0722
12	.0283	.0303	.0323	.0344	.0366	.0388	.0411	.0434	.0457	.0481
13	.0154	.0168	.0181	.0196	.0211	.0227	.0243	.0260	.0278	.0296
14	.0078	.0086	.0095	.0104	.0113	.0123	.0134	.0145	.0157	.0169
15	.0037	.0041	.0046	.0051	.0057	.0062	.0069	.0075	.0083	.0090
16	.0016	.0019	.0021	.0024	.0026	.0030	.0033	.0037	.0041	.0045
17	.0007	.0008	.0009	.0010	.0012	.0013	.0015	.0017	.0019	.0021
18	.0003	.0003	.0004	.0004	.0005	.0006	.0006	.0007	.0008	.0009
19	.0001	.0001	.0001	.0002	.0002	.0002	.0003	.0003	.0003	.0004
20	.0000	.0000	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0002
21	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001

x	λ 8.1	8.2	8.3	8.4	8.5	8.6	8.7	8.8	8.9	9.0
0	.0003	.0003	.0002	.0002	.0002	.0002	.0002	.0002	.0001	.0001
1	.0025	.0023	.0021	.0019	.0017	.0016	.0014	.0013	.0012	.0011
2	.0100	.0092	.0086	.0079	.0074	.0068	.0063	.0058	.0054	.0050
3	.0269	.0252	.0237	.0222	.0208	.0195	.0183	.0171	.0160	.0150
4	.0544	.0517	.0491	.0466	.0443	.0420	.0398	.0377	.0357	.0337
5	.0882	.0849	.0816	.0784	.0752	.0722	.0692	.0663	.0635	.0607
6	.1191	.1160	.1128	.1097	.1066	.1034	.1003	.0972	.0941	.0911
7	.1378	.1358	.1338	.1317	.1294	.1271	.1247	.1222	.1197	.1171
8	.1395	.1392	.1388	.1382	.1375	.1366	.1356	.1344	.1332	.1318
9	.1256	.1269	.1280	.1290	.1299	.1306	.1311	.1315	.1317	.1318
10	.1017	.1040	.1063	.1084	.1104	.1123	.1140	.1157	.1172	.1186
11	.0749	.0776	.0802	.0828	.0853	.0878	.0902	.0925	.0948	.0970
12	.0505	.0530	.0555	.0579	.0604	.0629	.0654	.0679	.0703	.0728
13	.0315	.0334	.0354	.0374	.0395	.0416	.0438	.0459	.0481	.0504
14	.0182	.0196	.0210	.0225	.0240	.0256	.0272	.0289	.0306	.0324
15	.0098	.0107	.0116	.0126	.0136	.0147	.0158	.0169	.0182	.0194
16	.0050	.0055	.0060	.0066	.0072	.0079	.0086	.0093	.0101	.0109
17	.0024	.0026	.0029	.0033	.0036	.0040	.0044	.0048	.0053	.0058
18	.0011	.0012	.0014	.0015	.0017	.0019	.0021	.0024	.0026	.0029
19	.0005	.0005	.0006	.0007	.0008	.0009	.0010	.0011	.0012	.0014
20	.0002	.0002	.0002	.0003	.0003	.0004	.0004	.0005	.0005	.0006
21	.0001	.0001	.0001	.0001	.0001	.0002	.0002	.0002	.0002	.0003
22	.0000	.0000	.0000	.0000	.0001	.0001	.0001	.0001	.0001	.0001

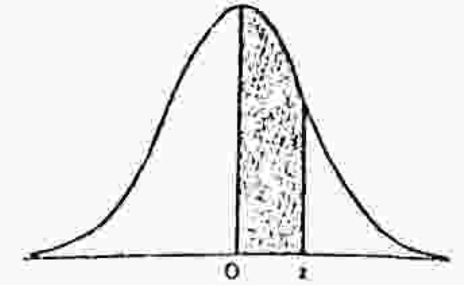
x	λ 9.1	9.2	9.3	9.4	9.5	9.6	9.7	9.8	9.9	10
0	[illegible]	[illegible]	[illegible]	.0001	.0001	.0001	.0001	.0001	.0001	.0000
1	[illegible]	[illegible]	[illegible]	.0008	.0007	.0007	.0006	.0005	.0005	.0005
2	[illegible]	[illegible]	[illegible]	.0037	.0034	.0031	.0029	.0027	.0025	.0023
3	[illegible]	[illegible]	[illegible]	.0115	.0107	.0100	.0093	.0087	.0081	.0076
4	[illegible]	[illegible]	[illegible]	.0269	.0254	.0240	.0226	.0213	.0201	.0189
5	[illegible]	[illegible]	[illegible]	.0506	.0483	.0460	.0439	.0418	.0398	.0378
6	[illegible]	[illegible]	[illegible]	.0793	.0764	.0736	.0709	.0682	.0656	.0631
7	[illegible]	[illegible]	[illegible]	.1064	.1037	.1010	.0982	.0955	.0928	.0901
8	[illegible]	[illegible]	[illegible]	.1251	.1232	.1212	.1191	.1170	.1148	.1126
9	[illegible]	[illegible]	[illegible]	.1306	.1300	.1293	.1284	.1274	.1263	.1251

تابع جدول (3)

x	λ = 9.1	9.2	9.3	9.4	9.5	9.6	9.7	9.8	9.9	10
10	.1198	.1210	.1219	.1228	.1235	.1241	.1245	.1249	.1250	.1251
11	.0991	.1012	.1031	.1049	.1067	.1083	.1098	.1112	.1125	.1137
12	.0752	.0776	.0799	.0822	.0844	.0866	.0888	.0908	.0928	.0948
13	.0526	.0549	.0572	.0594	.0617	.0640	.0662	.0685	.0707	.0729
14	.0342	.0361	.0380	.0399	.0419	.0439	.0459	.0479	.0500	.0521
15	.0208	.0221	.0235	.0250	.0265	.0281	.0297	.0313	.0330	.0347
16	.0118	.0127	.0137	.0147	.0157	.0168	.0180	.0192	.0204	.0217
17	.0063	.0069	.0075	.0081	.0088	.0095	.0103	.0111	.0119	.0128
18	.0032	.0035	.0039	.0042	.0046	.0051	.0055	.0060	.0065	.0071
19	.0015	.0017	.0019	.0021	.0023	.0026	.0028	.0031	.0034	.0037
20	.0007	.0008	.0009	.0010	.0011	.0012	.0014	.0015	.0017	.0019
21	.0003	.0003	.0004	.0004	.0005	.0006	.0006	.0007	.0008	.0009
22	.0001	.0001	.0002	.0002	.0002	.0002	.0003	.0003	.0004	.0004
23	.0000	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0002	.0002
24	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0001

x	λ = 11	12	13	14	15	16	17	18	19	20
0	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
1	.0002	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
2	.0010	.0004	.0002	.0001	.0000	.0000	.0000	.0000	.0000	.0000
3	.0037	.0018	.0008	.0004	.0002	.0001	.0000	.0000	.0000	.0000
4	.0102	.0053	.0027	.0013	.0006	.0003	.0001	.0001	.0000	.0000
5	.0224	.0127	.0070	.0037	.0019	.0010	.0005	.0002	.0001	.0001
6	.0411	.0255	.0152	.0087	.0048	.0026	.0014	.0007	.0004	.0002
7	.0646	.0437	.0281	.0174	.0104	.0060	.0034	.0018	.0010	.0005
8	.0888	.0655	.0457	.0304	.0194	.0120	.0072	.0042	.0024	.0013
9	.1085	.0874	.0661	.0473	.0324	.0213	.0135	.0083	.0050	.0029
10	.1194	.1048	.0859	.0663	.0486	.0341	.0230	.0150	.0095	.0058
11	.1194	.1144	.1015	.0844	.0663	.0496	.0355	.0245	.0164	.0106
12	.1094	.1144	.1099	.0984	.0829	.0661	.0504	.0368	.0259	.0176
13	.0926	.1056	.1099	.1060	.0956	.0814	.0658	.0509	.0378	.0271
14	.0728	.0905	.1021	.1060	.1024	.0930	.0800	.0655	.0514	.0387
15	.0534	.0724	.0885	.0989	.1024	.0992	.0906	.0786	.0650	.0516
16	.0367	.0543	.0719	.0866	.0960	.0992	.0963	.0884	.0772	.0646
17	.0237	.0383	.0550	.0713	.0847	.0934	.0963	.0936	.0863	.0760
18	.0145	.0255	.0397	.0554	.0706	.0830	.0909	.0936	.0911	.0844
19	.0084	.0161	.0272	.0409	.0557	.0699	.0814	.0887	.0911	.0888
20	.0046	.0097	.0177	.0286	.0418	.0559	.0692	.0798	.0866	.0888
21	.0024	.0055	.0109	.0191	.0299	.0426	.0560	.0684	.0783	.0846
22	.0012	.0030	.0065	.0121	.0204	.0310	.0433	.0560	.0676	.0769
23	.0006	.0016	.0037	.0074	.0133	.0216	.0320	.0438	.0559	.0669
24	.0003	.0008	.0020	.0043	.0083	.0144	.0227	.0328	.0442	.0557
25	.0001	.0004	.0010	.0024	.0050	.0092	.0154	.0237	.0336	.0446
26	.0000	.0002	.0005	.0013	.0029	.0057	.0101	.0164	.0246	.0343
27	.0000	.0001	.0002	.0007	.0016	.0034	.0063	.0109	.0173	.0254
28	.0000	.0000	.0001	.0003	.0009	.0019	.0038	.0070	.0117	.0181
29	.0000	.0000	.0001	.0002	.0004	.0011	.0023	.0044	.0077	.0125
30	.0000	.0000	.0000	.0001	.0002	.0006	.0013	.0026	.0049	.0083
31	.0000	.0000	.0000	.0000	.0001	.0003	.0007	.0015	.0030	.0054
32	.0000	.0000	.0000	.0000	.0001	.0001	.0004	.0009	.0018	.0034
33	.0000	.0000	.0000	.0000	.0000	.0001	.0002	.0005	.0010	.0020
34	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0002	.0006	.0012
35	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0003	.0007
36	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0002	.0004
37	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0002
38	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001
39	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001

جدول (4) : جدول التوزيع الطبيعي المعياري ، الذى يعطي القيم التى تنحصر بين 0 و z الموجبة ، أما القيم السالبة فيتم الحصول عليها بالتماثل .



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
0.7	.2580	.2611	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990
3.1	.4990	.4991	.4991	.4991	.4992	.4992	.4992	.4992	.4993	.4993
3.2	.4993	.4993	.4994	.4994	.4994	.4994	.4994	.4995	.4995	.4995
3.3	.4995	.4995	.4995	.4996	.4996	.4996	.4996	.4996	.4996	.4997
3.4	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4998
3.6	.4998	.4998	.4999	.4999	.4999	.4999	.4999	.4999	.4999	.4999
3.9	.5000									

المصدر :

Hoel and Jessen. Basic Statistics for Businness and Economics, 2nd ed., (1977) John Wiley & Sons.

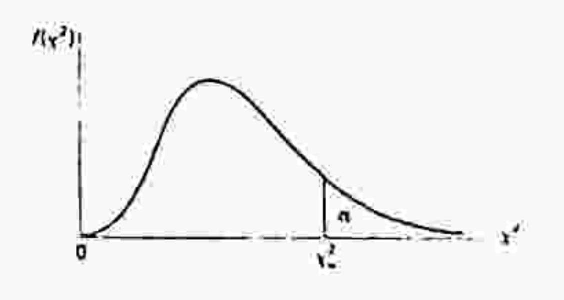
جدول (5) : القيم المئوية لتوزيع t ، حيث $P(t_{(n)} \geq t_{\alpha,n}) = \alpha$ ،

و n ترمز لدرجات الحرية .

n	α = .1	.05	.025	.01	.005
1	3.078	6.314	12.706	31.821	63.657
2	1.886	2.920	4.303	6.965	9.925
3	1.638	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
6	1.440	1.943	2.447	3.143	3.707
7	1.415	1.895	2.365	2.998	3.499
8	1.397	1.860	2.306	2.896	3.355
9	1.383	1.833	2.262	2.821	3.250
10	1.372	1.812	2.228	2.764	3.169
11	1.363	1.796	2.201	2.718	3.106
12	1.356	1.782	2.179	2.681	3.055
13	1.350	1.771	2.160	2.650	3.012
14	1.345	1.761	2.145	2.624	2.977
15	1.341	1.753	2.131	2.602	2.947
16	1.337	1.746	2.120	2.583	2.921
17	1.333	1.740	2.110	2.567	2.898
18	1.330	1.734	2.101	2.552	2.878
19	1.328	1.729	2.093	2.539	2.861
20	1.325	1.725	2.086	2.528	2.845
21	1.323	1.721	2.080	2.518	2.831
22	1.321	1.717	2.074	2.508	2.819
23	1.319	1.714	2.069	2.500	2.807
24	1.318	1.711	2.064	2.492	2.797
25	1.316	1.708	2.060	2.485	2.787
26	1.315	1.706	2.056	2.479	2.779
27	1.314	1.703	2.052	2.473	2.771
28	1.313	1.701	2.048	2.467	2.763
29	1.311	1.699	2.045	2.462	2.756
30	1.310	1.697	2.042	2.457	2.750
40	1.303	1.684	2.021	2.423	2.704
60	1.296	1.671	2.000	2.390	2.660
120	1.290	1.661	1.984	2.358	2.626
∞	1.282	1.645	1.960	2.326	2.576

المصدر : Dunn abd Clark , Applied Statistics , (1974) , John Wiley & Sons .

جدول (6) : القيم المئوية لتوزيع مربع كاى ، حيث $P(\chi^2_{(n)} \geq \chi^2_{\alpha,n}) = \alpha$ ،
و n ترمز لدرجات الحرية .



n \ α	.99	.975	.95	.90	.80	.70	.50	.30	.20	.10	.05	.025	.01	.001
1	.000157	0.00098	.00393	.0158	.0642	.148	.455	1.074	1.642	2.706	3.841	5.0238	6.635	10.827
2	.0201	0.0506	.103	.211	.446	.713	1.386	2.408	3.219	4.605	5.991	7.3780	9.210	13.815
3	.115	0.216	.352	.584	1.005	1.424	2.366	3.665	4.642	6.251	7.815	9.348	11.345	16.266
4	.297	0.484	.711	1.064	1.649	2.195	3.357	4.878	5.989	7.779	9.488	11.143	13.277	18.467
5	.554	0.831	1.145	1.610	2.343	3.000	4.351	6.064	7.289	9.236	11.070	12.832	15.086	20.515
6	.872	1.237	1.635	2.204	3.070	3.828	5.348	7.231	8.558	10.645	12.592	14.449	16.812	22.457
7	1.239	1.690	2.167	2.833	3.822	4.671	6.346	8.383	9.803	12.017	14.067	16.013	18.475	24.322
8	1.646	2.180	2.733	3.490	4.594	5.527	7.344	9.524	11.030	13.362	15.507	17.535	20.090	26.125
9	2.088	2.700	3.325	4.168	5.380	6.393	8.343	10.656	12.242	14.684	16.919	19.023	21.666	27.877
10	2.558	3.247	3.940	4.865	6.179	7.267	9.342	11.781	13.442	15.987	18.307	20.483	23.209	29.588
11	3.053	3.816	4.575	5.578	6.989	8.148	10.341	12.899	14.631	17.275	19.675	21.920	24.725	31.264
12	3.571	4.404	5.226	6.304	7.807	9.034	11.340	14.011	15.812	18.549	21.026	23.337	26.217	32.909
13	4.107	5.009	5.892	7.042	8.634	9.926	12.340	15.119	16.985	19.812	22.362	24.736	27.688	34.528
14	4.660	5.629	6.571	7.790	9.467	10.821	13.339	16.222	18.151	21.064	23.685	26.119	29.141	36.123
15	5.229	6.262	7.261	8.547	10.307	11.721	14.339	17.322	19.311	22.307	24.996	27.488	30.578	37.697

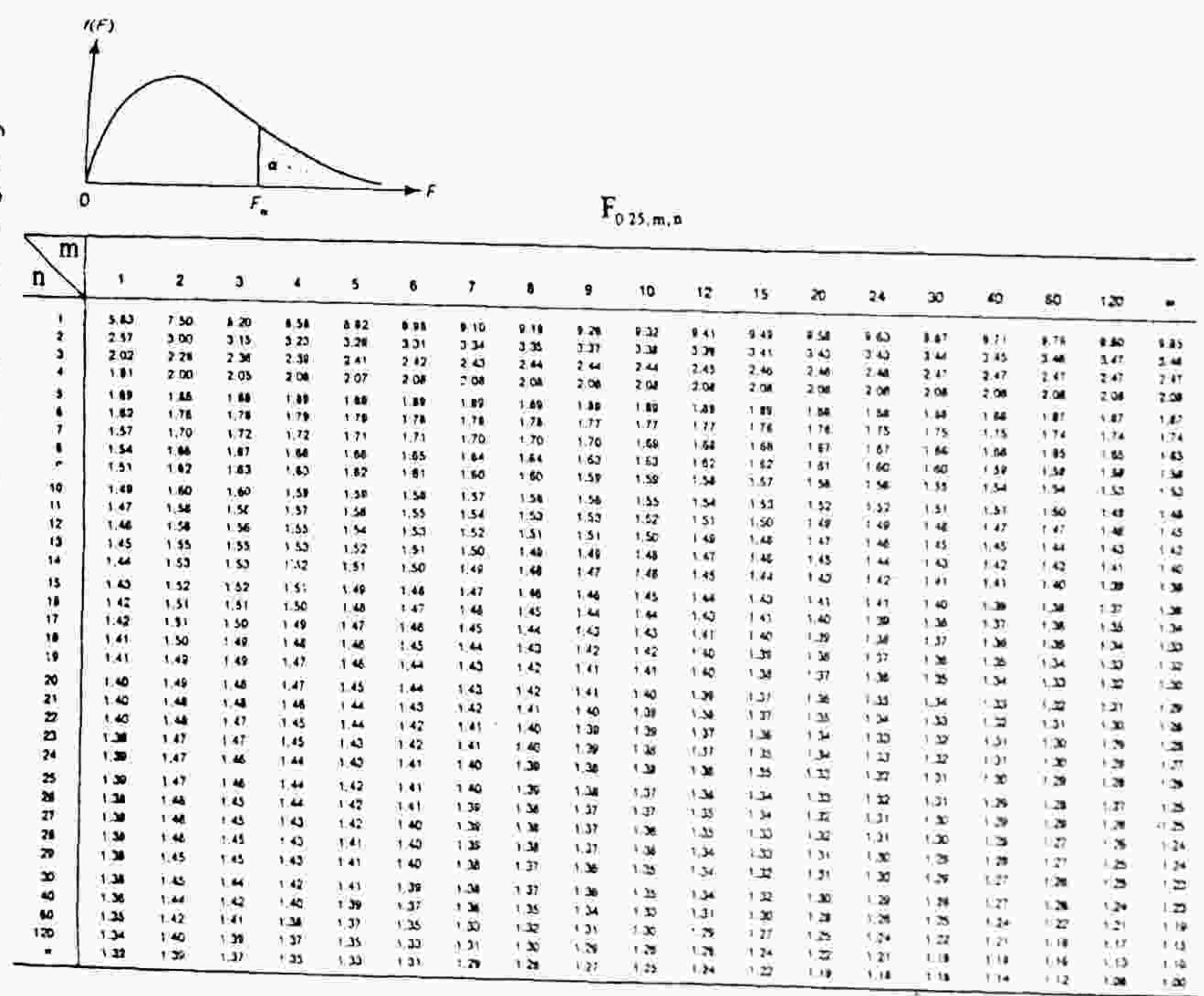
المصدر : . Biometrika Tables for Statisticions , vol. II , 1972

تابع جدول (6)

n \ P	.99	.975	.95	.90	.80	.70	.50	.30	.20	.10	.05	.025	.01	.001
16	5.812	6.908	7.962	9.312	11.152	12.624	15.338	18.418	20.465	23.542	26.296	28.845	32.000	39.252
17	6.408	7.564	8.672	10.085	12.002	13.531	16.338	19.511	21.615	24.769	27.587	30.191	33.409	40.790
18	7.015	8.231	9.390	10.865	12.857	14.440	17.338	20.601	22.760	25.989	28.869	31.526	34.805	42.312
19	7.633	8.907	10.117	11.651	13.716	15.352	18.338	21.689	23.900	27.204	30.144	32.852	36.191	43.820
20	8.260	9.591	10.851	12.443	14.578	16.266	19.337	22.775	25.038	28.412	31.410	34.170	37.566	45.315
21	8.897	10.283	11.591	13.240	15.445	17.182	20.337	23.858	26.171	29.615	32.671	35.479	38.932	46.797
22	9.542	10.982	12.338	14.041	16.314	18.101	21.337	24.939	27.301	30.813	33.924	36.781	40.289	48.268
23	10.196	11.689	13.091	14.848	17.187	19.021	22.337	26.018	28.429	32.007	35.172	38.076	41.638	49.728
24	10.856	12.401	13.848	15.659	18.062	19.943	23.337	27.096	29.553	33.196	36.415	39.364	42.980	51.179
25	11.524	13.120	14.611	16.473	18.940	20.867	24.337	28.172	30.675	34.382	37.652	40.646	44.314	52.620
26	12.198	13.844	15.379	17.292	19.820	21.792	25.336	29.246	31.795	35.563	38.885	41.923	45.642	54.052
27	12.879	14.573	16.151	18.114	20.703	22.719	26.336	30.319	32.912	36.741	40.113	43.194	46.963	55.476
28	13.565	15.308	16.928	18.939	21.588	23.647	27.336	31.391	34.027	37.916	41.337	44.461	48.278	56.893
29	14.256	16.047	17.708	19.768	22.475	24.577	28.336	32.461	35.139	39.087	42.557	45.722	49.588	58.302
30	14.953	16.791	18.493	20.599	23.364	25.508	29.336	33.530	36.250	40.256	43.773	46.979	50.892	59.703

جدول (7) : القيم المئوية لتوزيع F ، حيث $P(F_{m,n} \geq F_{\alpha,m,n}) = \alpha$ ،

و m ترمز لدرجات حرية البسط و n ترمز لدرجات حرية المقام .



$F_{0.25,m,n}$

n \ m	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	5.83	7.50	8.20	8.58	8.82	8.98	9.10	9.19	9.26	9.32	9.41	9.49	9.58	9.63	9.67	9.71	9.76	9.80	9.85
2	2.57	3.00	3.15	3.23	3.28	3.31	3.34	3.35	3.37	3.38	3.39	3.41	3.43	3.43	3.44	3.45	3.46	3.47	3.48
3	2.02	2.28	2.36	2.39	2.41	2.42	2.43	2.44	2.44	2.44	2.45	2.46	2.46	2.46	2.47	2.47	2.47	2.47	2.47
4	1.81	2.00	2.05	2.06	2.07	2.08	2.08	2.08	2.08	2.08	2.08	2.08	2.08	2.08	2.08	2.08	2.08	2.08	2.08
5	1.69	1.85	1.88	1.89	1.89	1.89	1.89	1.89	1.89	1.89	1.89	1.89	1.88	1.88	1.88	1.88	1.87	1.87	1.87
6	1.62	1.76	1.78	1.79	1.79	1.78	1.78	1.78	1.77	1.77	1.77	1.76	1.76	1.75	1.75	1.75	1.74	1.74	1.74
7	1.57	1.70	1.72	1.72	1.71	1.71	1.70	1.70	1.69	1.69	1.68	1.68	1.67	1.67	1.66	1.66	1.65	1.65	1.65
8	1.54	1.66	1.67	1.66	1.66	1.65	1.64	1.64	1.63	1.63	1.62	1.62	1.61	1.60	1.60	1.59	1.59	1.58	1.58
9	1.51	1.62	1.63	1.63	1.62	1.61	1.60	1.60	1.59	1.59	1.58	1.57	1.56	1.56	1.55	1.54	1.54	1.53	1.53
10	1.49	1.60	1.60	1.59	1.59	1.58	1.57	1.56	1.56	1.55	1.54	1.53	1.52	1.52	1.51	1.51	1.50	1.49	1.48
11	1.47	1.58	1.58	1.57	1.56	1.55	1.54	1.53	1.53	1.52	1.51	1.50	1.49	1.49	1.48	1.47	1.47	1.46	1.45
12	1.46	1.56	1.56	1.55	1.54	1.53	1.52	1.51	1.51	1.50	1.49	1.48	1.47	1.46	1.45	1.45	1.44	1.43	1.42
13	1.45	1.55	1.55	1.53	1.52	1.51	1.50	1.49	1.49	1.48	1.47	1.46	1.45	1.44	1.43	1.42	1.42	1.41	1.40
14	1.44	1.53	1.53	1.52	1.51	1.50	1.49	1.48	1.47	1.46	1.45	1.44	1.43	1.42	1.41	1.41	1.40	1.39	1.38
15	1.43	1.52	1.52	1.51	1.49	1.48	1.47	1.46	1.46	1.45	1.44	1.43	1.41	1.41	1.40	1.39	1.38	1.37	1.36
16	1.42	1.51	1.51	1.50	1.48	1.47	1.46	1.45	1.44	1.44	1.43	1.41	1.40	1.39	1.38	1.37	1.36	1.35	1.34
17	1.42	1.51	1.50	1.49	1.47	1.46	1.45	1.44	1.43	1.43	1.41	1.40	1.39	1.38	1.37	1.36	1.35	1.34	1.33
18	1.41	1.50	1.49	1.48	1.46	1.45	1.44	1.43	1.42	1.42	1.40	1.39	1.38	1.37	1.36	1.35	1.34	1.33	1.32
19	1.41	1.49	1.49	1.47	1.46	1.44	1.43	1.42	1.41	1.41	1.40	1.38	1.37	1.36	1.35	1.34	1.33	1.32	1.30
20	1.40	1.49	1.48	1.47	1.45	1.44	1.43	1.42	1.41	1.40	1.39	1.37	1.36	1.35	1.34	1.33	1.32	1.31	1.29
21	1.40	1.48	1.48	1.46	1.44	1.43	1.42	1.41	1.40	1.39	1.38	1.37	1.35	1.34	1.33	1.32	1.31	1.30	1.28
22	1.40	1.48	1.47	1.45	1.44	1.42	1.41	1.40	1.39	1.39	1.37	1.36	1.34	1.33	1.32	1.31	1.30	1.29	1.28
23	1.39	1.47	1.47	1.45	1.43	1.42	1.41	1.40	1.39	1.38	1.37	1.35	1.34	1.33	1.32	1.31	1.30	1.28	1.27
24	1.39	1.47	1.46	1.44	1.43	1.41	1.40	1.39	1.38	1.38	1.36	1.35	1.33	1.32	1.31	1.30	1.29	1.28	1.26
25	1.39	1.47	1.46	1.44	1.42	1.41	1.40	1.39	1.38	1.37	1.36	1.34	1.33	1.32	1.31	1.29	1.28	1.27	1.25
26	1.38	1.46	1.45	1.44	1.42	1.41	1.39	1.38	1.37	1.37	1.35	1.34	1.32	1.31	1.30	1.29	1.28	1.26	1.25
27	1.38	1.46	1.45	1.43	1.42	1.40	1.39	1.38	1.37	1.36	1.35	1.33	1.32	1.31	1.30	1.28	1.27	1.26	1.24
28	1.38	1.46	1.45	1.43	1.41	1.40	1.39	1.38	1.37	1.36	1.34	1.33	1.31	1.30	1.29	1.28	1.27	1.25	1.24
29	1.38	1.45	1.45	1.43	1.41	1.40	1.38	1.37	1.36	1.35	1.34	1.32	1.31	1.30	1.29	1.27	1.26	1.25	1.23
30	1.38	1.45	1.44	1.42	1.41	1.39	1.38	1.37	1.36	1.35	1.34	1.32	1.30	1.29	1.28	1.27	1.26	1.24	1.23
40	1.36	1.44	1.42	1.40	1.39	1.37	1.36	1.35	1.34	1.33	1.31	1.30	1.28	1.26	1.25	1.24	1.22	1.21	1.19
60	1.35	1.42	1.41	1.38	1.37	1.35	1.33	1.32	1.31	1.30	1.29	1.27	1.25	1.24	1.22	1.21	1.19	1.17	1.15
120	1.34	1.40	1.39	1.37	1.35	1.33	1.31	1.30	1.29	1.28	1.26	1.24	1.22	1.21	1.19	1.18	1.16	1.13	1.10
∞	1.32	1.39	1.37	1.35	1.33	1.31	1.29	1.28	1.27	1.25	1.24	1.22	1.19	1.18	1.16	1.14	1.12	1.08	1.00

المصدر : Biometrika Tables for Statistics , vol. 1 , 1966 .

تابع جدول (7)

$F_{0.10,m,n}$

n \ m	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	39.86	49.50	53.59	55.83	57.24	58.20	58.91	59.44	59.86	60.19	60.71	61.22	61.74	62.00	62.26	62.53	62.79	63.06	63.33
2	8.53	9.00	9.16	9.24	9.29	9.33	9.35	9.37	9.38	9.39	9.41	9.42	9.44	9.45	9.46	9.47	9.47	9.48	9.49
3	5.54	5.46	5.39	5.34	5.31	5.28	5.27	5.25	5.24	5.23	5.22	5.20	5.18	5.18	5.17	5.16	5.15	5.14	5.13
4	4.54	4.32	4.19	4.11	4.05	4.01	3.98	3.95	3.94	3.92	3.90	3.87	3.84	3.83	3.82	3.80	3.79	3.78	3.76
5	4.06	3.78	3.62	3.52	3.45	3.40	3.37	3.34	3.32	3.30	3.27	3.24	3.21	3.19	3.17	3.16	3.14	3.12	3.10
6	3.78	3.46	3.29	3.18	3.11	3.05	3.01	2.98	2.96	2.94	2.90	2.87	2.84	2.82	2.80	2.78	2.76	2.74	2.72
7	3.59	3.26	3.07	2.96	2.88	2.83	2.78	2.75	2.72	2.70	2.67	2.63	2.59	2.58	2.56	2.54	2.51	2.49	2.47
8	3.46	3.11	2.92	2.81	2.73	2.67	2.62	2.59	2.56	2.54	2.50	2.46	2.42	2.40	2.38	2.36	2.34	2.32	2.29
9	3.36	3.01	2.81	2.69	2.61	2.55	2.51	2.47	2.44	2.42	2.38	2.34	2.30	2.28	2.25	2.23	2.21	2.18	2.16
10	3.29	2.92	2.73	2.61	2.52	2.46	2.41	2.38	2.35	2.32	2.28	2.24	2.20	2.18	2.16	2.13	2.11	2.08	2.06
11	3.23	2.86	2.66	2.54	2.45	2.39	2.34	2.30	2.27	2.25	2.21	2.17	2.12	2.10	2.08	2.05	2.03	2.00	1.97
12	3.18	2.81	2.61	2.48	2.39	2.33	2.28	2.24	2.21	2.19	2.15	2.10	2.06	2.04	2.01	1.99	1.96	1.93	1.90
13	3.14	2.76	2.56	2.43	2.35	2.28	2.23	2.20	2.16	2.14	2.10	2.05	2.01	1.98	1.96	1.93	1.90	1.88	1.85
14	3.10	2.73	2.52	2.39	2.31	2.24	2.19	2.15	2.12	2.10	2.05	2.01	1.96	1.94	1.91	1.89	1.86	1.83	1.80
15	3.07	2.70	2.49	2.36	2.27	2.21	2.16	2.12	2.09	2.06	2.02	1.97	1.92	1.90	1.87	1.85	1.82	1.79	1.76
16	3.05	2.67	2.46	2.33	2.24	2.18	2.13	2.09	2.06	2.03	1.99	1.94	1.89	1.87	1.84	1.81	1.78	1.75	1.72
17	3.03	2.64	2.44	2.31	2.22	2.15	2.10	2.06	2.03	2.00	1.96	1.91	1.86	1.84	1.81	1.78	1.75	1.72	1.69
18	3.01	2.62	2.42	2.29	2.20	2.13	2.08	2.04	2.00	1.98	1.93	1.89	1.84	1.81	1.78	1.75	1.72	1.69	1.66
19	2.99	2.61	2.40	2.27	2.18	2.11	2.06	2.02	1.98	1.96	1.91	1.86	1.81	1.79	1.76	1.73	1.70	1.67	1.63
20	2.97	2.59	2.38	2.25	2.16	2.09	2.04	2.00	1.96	1.94	1.89	1.84	1.79	1.77	1.74	1.71	1.68	1.64	1.61
21	2.96	2.57	2.36	2.23	2.14	2.08	2.02	1.98	1.95	1.92	1.87	1.83	1.78	1.75	1.72	1.69	1.66	1.62	1.59
22	2.95	2.56	2.35	2.22	2.13	2.06	2.01	1.97	1.93	1.90	1.86	1.81	1.76	1.73	1.70	1.67	1.64	1.60	1.57
23	2.94	2.55	2.34	2.21	2.11	2.05	1.99	1.95	1.92	1.89	1.84	1.80	1.74	1.72	1.69	1.66	1.62	1.59	1.55
24	2.93	2.54	2.33	2.19	2.10	2.04	1.98	1.94	1.91	1.88	1.83	1.78	1.73	1.70	1.67	1.64	1.61	1.57	1.53
25	2.92	2.53	2.32	2.18	2.09	2.02	1.97	1.93	1.89	1.87	1.82	1.77	1.72	1.69	1.66	1.63	1.59	1.56	1.52
26	2.91	2.52	2.31	2.17	2.08	2.01	1.96	1.92	1.88	1.86	1.81	1.76	1.71	1.68	1.65	1.61	1.58	1.54	1.50
27	2.90	2.51	2.30	2.17	2.07	2.00	1.95	1.91	1.87	1.85	1.80	1.75	1.70	1.67	1.64	1.60	1.57	1.53	1.49
28	2.89	2.50	2.29	2.16	2.06	2.00	1.94	1.90	1.87	1.84	1.79	1.74	1.69	1.66	1.63	1.59	1.56	1.52	1.48
29	2.89	2.50	2.28	2.15	2.06	1.99	1.93	1.89	1.86	1.83	1.78	1.73	1.68	1.65	1.62	1.58	1.55	1.51	1.47
30	2.88	2.49	2.28	2.14	2.05	1.98	1.93	1.88	1.85	1.82	1.77	1.72	1.67	1.64	1.61	1.57	1.54	1.50	1.46
40	2.84	2.44	2.23	2.09	2.00	1.93	1.87	1.83	1.79	1.76	1.71	1.66	1.61	1.57	1.54	1.51	1.47	1.42	1.38
60	2.79	2.39	2.18	2.04	1.95	1.87	1.82	1.77	1.74	1.71	1.66	1.60	1.54	1.51	1.48	1.44	1.40	1.35	1.29
120	2.75	2.35	2.13	1.99	1.90	1.82	1.77	1.72	1.68	1.65	1.60	1.55	1.48	1.45	1.41	1.37	1.32	1.26	1.19
∞	2.71	2.30	2.08	1.94	1.85	1.77	1.72	1.67	1.63	1.60	1.55	1.49	1.42	1.38	1.34	1.30	1.24	1.17	1.00

$F_{0.05,m,n}$

n \ m	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	161.4	199.5	215.7	224.6	230.2	234.0	236.8	238.9	240.5	241.9	243.9	245.9	248.0	249.1	250.1	251.1	252.2	253.3	254.3
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.41	19.43	19.45	19.45	19.46	19.47	19.48	19.49	19.50
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.70	8.66	8.64	8.62	8.59	8.57	8.55	8.53
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.86	5.80	5.77	5.75	5.72	5.69	5.66	5.63
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.53	4.50	4.46	4.43	4.40	4.36
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.94	3.87	3.84	3.81	3.77	3.74	3.70	3.67
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.51	3.44	3.41	3.38	3.34	3.30	3.27	3.23
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.22	3.15	3.12	3.08	3.04	3.01	2.97	2.93
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01	2.94	2.90	2.86	2.83	2.79	2.75	2.71
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.85	2.77	2.74	2.70	2.66	2.62	2.58	2.54
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.79	2.72	2.65	2.61	2.57	2.53	2.49	2.45	2.40
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.62	2.54	2.51	2.47	2.43	2.38	2.34	2.30
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.60	2.53	2.46	2.42	2.38	2.34	2.30	2.25	2.21
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53	2.46	2.39	2.35	2.31	2.27	2.22	2.18	2.13
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.40	2.33	2.29	2.25	2.20	2.16	2.11	2.07
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.35	2.28	2.24	2.19	2.15	2.11	2.06	2.01
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.38	2.31	2.23	2.19	2.15	2.10	2.06	2.01	1.96
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.27	2.19	2.15	2.11	2.06	2.02	1.97	1.92
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.93	1.88
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.20	2.12	2.08	2.04	1.99	1.95	1.90	1.84
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.25	2.18	2.10	2.05	2.01	1.96	1.92	1.87	1.81
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.23	2.15	2.07	2.03	1.98	1.94	1.89	1.84	1.78
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2.20	2.13	2.05	2.01	1.96	1.91	1.86	1.81	1.76
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.18	2.11	2.03	1.98	1.94	1.89	1.84	1.79	1.73
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.16	2.09	2.01	1.96	1.92	1.87	1.82	1.77	1.71
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.15	2.07	1.99	1.95	1.90	1.85	1.80	1.75	1.69
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.20	2.13	2.06	1.97	1.93	1.88	1.84	1.79	1.73	1.67
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19	2.12	2.04	1.96	1.91	1.87	1.82	1.77	1.71	1.65
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22	2.18	2.10	2.03	1.94	1.90	1.85	1.81	1.75	1.70	1.64
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09	2.01	1.93	1.89	1.84	1.79	1.74	1.68	1.62
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.00	1.92	1.84	1.79	1.74	1.69	1.64	1.58	1.51
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.92	1.84	1.75	1.70	1.65	1.59	1.53	1.47	1.39
120	3.92	3.07	2.68	2.45	2.29	2.17	2.09	2.02	1.96	1.91	1.83	1.75	1.66	1.61	1.55	1.50	1.43	1.35	1.25
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.75	1.67	1.57	1.52	1.46	1.39	1.32	1.22	1.00

تابع جدول (7)

$F_{0.025,m,n}$

n \ m	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	647.8	799.5	864.2	899.6	921.8	937.1	948.2	956.7	963.3	968.6	976.7	984.9	993.1	997.2	1001	1006	1010	1014	1018
2	38.51	39.00	39.17	39.25	39.30	39.33	39.36	39.37	39.39	39.40	39.41	39.43	39.45	39.46	39.46	39.47	39.48	39.49	39.50
3	17.44	16.04	15.44	15.10	14.88	14.73	14.62	14.54	14.47	14.42	14.34	14.25	14.17	14.12	14.08	14.04	13.99	13.95	13.90
4	12.22	10.65	9.98	9.60	9.36	9.20	9.07	8.98	8.90	8.84	8.75	8.66	8.56	8.51	8.46	8.41	8.36	8.31	8.26
5	10.01	8.43	7.76	7.39	7.15	6.98	6.85	6.76	6.68	6.62	6.52	6.43	6.33	6.28	6.23	6.18	6.12	6.07	6.02
6	8.81	7.26	6.60	6.23	5.99	5.82	5.70	5.60	5.52	5.46	5.37	5.27	5.17	5.12	5.07	5.01	4.96	4.90	4.85
7	8.07	6.54	5.89	5.52	5.29	5.12	4.99	4.90	4.82	4.76	4.67	4.57	4.47	4.42	4.36	4.31	4.25	4.20	4.14
8	7.57	6.06	5.42	5.05	4.82	4.65	4.53	4.43	4.36	4.30	4.20	4.10	4.00	3.95	3.89	3.84	3.78	3.73	3.67
9	7.21	5.71	5.08	4.72	4.48	4.32	4.20	4.10	4.03	3.96	3.87	3.77	3.67	3.61	3.56	3.51	3.45	3.39	3.33
10	6.94	5.46	4.83	4.47	4.24	4.07	3.95	3.85	3.78	3.72	3.62	3.52	3.42	3.37	3.31	3.26	3.20	3.14	3.08
11	6.72	5.26	4.63	4.28	4.04	3.88	3.76	3.66	3.59	3.53	3.43	3.33	3.23	3.17	3.12	3.06	3.00	2.94	2.88
12	6.55	5.10	4.47	4.12	3.89	3.73	3.61	3.51	3.44	3.37	3.28	3.18	3.07	3.02	2.96	2.91	2.85	2.79	2.72
13	6.41	4.97	4.35	4.00	3.77	3.60	3.48	3.39	3.31	3.25	3.15	3.05	2.95	2.89	2.84	2.78	2.72	2.66	2.60
14	6.30	4.86	4.24	3.89	3.66	3.50	3.38	3.29	3.21	3.15	3.05	2.95	2.84	2.79	2.73	2.67	2.61	2.55	2.49
15	6.20	4.77	4.15	3.80	3.58	3.41	3.29	3.20	3.12	3.06	2.96	2.86	2.76	2.70	2.64	2.59	2.52	2.46	2.40
16	6.12	4.69	4.08	3.73	3.50	3.34	3.22	3.12	3.05	2.99	2.89	2.79	2.68	2.63	2.57	2.51	2.45	2.38	2.32
17	6.04	4.62	4.01	3.66	3.44	3.28	3.16	3.06	2.98	2.92	2.82	2.72	2.62	2.56	2.50	2.44	2.38	2.32	2.25
18	5.98	4.56	3.95	3.61	3.38	3.22	3.10	3.01	2.93	2.87	2.77	2.67	2.56	2.50	2.44	2.38	2.32	2.26	2.19
19	5.92	4.51	3.90	3.56	3.33	3.17	3.05	2.96	2.88	2.82	2.72	2.62	2.51	2.45	2.39	2.33	2.27	2.20	2.13
20	5.87	4.46	3.86	3.51	3.29	3.13	3.01	2.91	2.84	2.77	2.68	2.57	2.46	2.41	2.35	2.29	2.22	2.16	2.09
21	5.83	4.42	3.82	3.48	3.25	3.09	2.97	2.87	2.80	2.73	2.64	2.53	2.42	2.37	2.31	2.25	2.18	2.11	2.04
22	5.79	4.38	3.78	3.44	3.22	3.05	2.93	2.84	2.76	2.70	2.60	2.50	2.39	2.33	2.27	2.21	2.14	2.08	2.00
23	5.75	4.35	3.75	3.41	3.18	3.02	2.90	2.81	2.73	2.67	2.57	2.47	2.36	2.30	2.24	2.18	2.11	2.04	1.97
24	5.72	4.32	3.72	3.38	3.15	2.99	2.87	2.78	2.70	2.64	2.54	2.44	2.33	2.27	2.21	2.15	2.08	2.01	1.94
25	5.69	4.29	3.69	3.35	3.13	2.97	2.85	2.75	2.68	2.61	2.51	2.41	2.30	2.24	2.18	2.12	2.05	1.98	1.91
26	5.66	4.27	3.67	3.33	3.10	2.94	2.82	2.73	2.65	2.59	2.49	2.39	2.28	2.22	2.16	2.09	2.03	1.95	1.88
27	5.63	4.24	3.65	3.31	3.08	2.92	2.80	2.71	2.63	2.57	2.47	2.36	2.25	2.19	2.13	2.07	2.00	1.93	1.85
28	5.61	4.22	3.63	3.29	3.06	2.90	2.78	2.69	2.61	2.55	2.45	2.34	2.23	2.17	2.11	2.05	1.98	1.91	1.83
29	5.59	4.20	3.61	3.27	3.04	2.88	2.76	2.67	2.59	2.53	2.43	2.32	2.21	2.15	2.09	2.03	1.96	1.89	1.81
30	5.57	4.18	3.59	3.25	3.03	2.87	2.75	2.65	2.57	2.51	2.41	2.31	2.20	2.14	2.07	2.01	1.94	1.87	1.79
40	5.42	4.05	3.46	3.13	2.90	2.74	2.62	2.53	2.45	2.39	2.29	2.18	2.07	2.01	1.94	1.88	1.80	1.72	1.64
60	5.29	3.93	3.34	3.01	2.79	2.63	2.51	2.41	2.33	2.27	2.17	2.06	1.94	1.88	1.82	1.74	1.67	1.58	1.48
120	5.15	3.80	3.23	2.89	2.67	2.52	2.39	2.30	2.22	2.16	2.05	1.94	1.82	1.76	1.69	1.61	1.53	1.43	1.31
∞	5.02	3.69	3.12	2.79	2.57	2.41	2.29	2.19	2.11	2.05	1.94	1.83	1.71	1.64	1.57	1.48	1.39	1.27	1.00

تابع جدول (7)

$F_{0.01,m,n}$

n \ m	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	4052	4999.5	5403	5625	5764	5859	5928	5982	6022	6056	6106	6157	6209	6235	6261	6287	6313	6339	6366
2	98.50	99.00	99.17	99.25	99.30	99.33	99.36	99.37	99.39	99.40	99.42	99.43	99.45	99.46	99.47	99.47	99.48	99.49	99.50
3	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.35	27.23	27.05	26.87	26.69	26.60	26.50	26.41	26.32	26.22	26.13
4	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66	14.55	14.37	14.20	14.02	13.93	13.84	13.75	13.65	13.56	13.46
5	16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16	10.05	9.89	9.72	9.55	9.47	9.38	9.29	9.20	9.11	9.02
6	13.75	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87	7.72	7.56	7.40	7.31	7.23	7.14	7.06	6.97	6.88
7	12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72	6.62	6.47	6.31	6.16	6.07	5.99	5.91	5.82	5.74	5.65
8	11.26	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91	5.81	5.67	5.52	5.36	5.28	5.20	5.12	5.03	4.95	4.86
9	10.56	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35	5.26	5.11	4.96	4.81	4.73	4.65	4.57	4.48	4.40	4.31
10	10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85	4.71	4.56	4.41	4.33	4.25	4.17	4.08	4.00	3.91
11	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63	4.54	4.40	4.25	4.10	4.02	3.94	3.86	3.78	3.69	3.60
12	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39	4.30	4.16	4.01	3.86	3.78	3.70	3.62	3.54	3.45	3.36
13	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19	4.10	3.96	3.82	3.66	3.59	3.51	3.43	3.34	3.25	3.17
14	8.86	6.51	5.56	5.04	4.69	4.46	4.28	4.14	4.03	3.94	3.80	3.66	3.51	3.43	3.35	3.27	3.18	3.09	3.00
15	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89	3.80	3.67	3.52	3.37	3.29	3.21	3.13	3.05	2.96	2.87
16	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78	3.69	3.55	3.41	3.26	3.18	3.10	3.02	2.93	2.84	2.75
17	8.40	6.11	5.18	4.67	4.34	4.10	3.93	3.79	3.68	3.59	3.46	3.31	3.16	3.08	3.00	2.92	2.83	2.75	2.65
18	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60	3.51	3.37	3.23	3.08	3.00	2.92	2.84	2.75	2.66	2.57
19	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52	3.43	3.30	3.15	3.00	2.92	2.84	2.76	2.67	2.58	2.49
20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46	3.37	3.23	3.09	2.94	2.86	2.78	2.69	2.61	2.52	2.42
21	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40	3.31	3.17	3.03	2.88	2.80	2.72	2.64	2.55	2.46	2.36
22	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35	3.26	3.12	2.98	2.83	2.75	2.67	2.58	2.50	2.40	2.31
23	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30	3.21	3.07	2.93	2.78	2.70	2.62	2.54	2.45	2.35	2.26
24	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26	3.17	3.03	2.89	2.74	2.66	2.58	2.49	2.40	2.31	2.21
25	7.77	5.57	4.68	4.18	3.85	3.63	3.46	3.32	3.22	3.13	2.99	2.85	2.70	2.62	2.54	2.45	2.36	2.27	2.17
26	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.18	3.09	2.96	2.81	2.66	2.58	2.50	2.42	2.33	2.23	2.13
27	7.68	5.49	4.60	4.11	3.78	3.56	3.39	3.26	3.15	3.06	2.93	2.78	2.63	2.55	2.47	2.38	2.29	2.20	2.10
28	7.64	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.12	3.03	2.90	2.75	2.60	2.52	2.44	2.35	2.26	2.17	2.06
29	7.60	5.42	4.54	4.04	3.73	3.50	3.33	3.20	3.09	3.00	2.87	2.73	2.57	2.49	2.41	2.33	2.23	2.14	2.03
30	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07	2.98	2.84	2.70	2.55	2.47	2.39	2.30	2.21	2.11	2.01
40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89	2.80	2.66	2.52	2.37	2.29	2.20	2.11	2.02	1.92	1.80
60	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63	2.50	2.35	2.20	2.12	2.03	1.94	1.84	1.73	1.60
120	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56	2.47	2.34	2.19	2.03	1.95	1.86	1.76	1.66	1.53	1.38
∞	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41	2.32	2.18	2.04	1.88	1.79	1.70	1.59	1.47	1.32	1.00

جدول (8) : مئويات اختبار مان - وايتني .

n_1	p	$n_2=2$	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
2	.001	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	.005	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
	.01	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	2	2
	.025	0	0	0	0	0	0	1	1	1	1	2	2	2	2	2	3	3	3	3
	.05	0	0	0	1	1	1	2	2	2	2	3	3	4	4	4	4	5	5	5
	.10	0	1	1	2	2	2	3	3	4	4	5	5	5	6	6	7	7	8	8
3	.001	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1
	.005	0	0	0	0	0	0	0	1	1	1	2	2	2	3	3	3	3	4	4
	.01	0	0	0	0	0	1	1	2	2	2	3	3	3	4	4	5	5	5	6
	.025	0	0	0	1	2	2	3	3	4	4	5	5	6	6	7	7	8	8	9
	.05	0	1	1	2	3	3	4	5	5	6	6	7	8	8	9	10	10	11	12
	.10	1	2	2	3	4	5	6	6	7	8	9	10	11	11	12	13	14	15	16
4	.001	0	0	0	0	0	0	0	0	1	1	1	2	2	2	3	3	4	4	4
	.005	0	0	0	0	1	1	2	2	3	3	4	4	5	6	6	7	7	8	9
	.01	0	0	0	1	2	2	3	4	4	5	6	6	7	9	8	9	10	10	11
	.025	0	0	1	2	3	4	5	5	6	7	8	9	10	11	12	12	13	14	15
	.05	0	1	2	3	4	5	6	7	8	9	10	11	12	13	15	16	17	18	19
	.10	1	2	4	5	6	7	8	10	11	12	13	14	16	17	18	19	21	22	23

المصدر :

Extended Tables of Critical Values for Wilcoxon's Test Statistic, Biometrika, 50 (1963).

تابع جدول (8)

n_1	p	$n_2=2$	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
5	.001	0	0	0	0	0	0	1	2	2	3	3	4	4	5	6	6	7	8	8
	.005	0	0	0	1	2	2	3	4	5	6	7	8	8	9	10	11	12	13	14
	.01	0	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
	.025	0	1	2	3	4	6	7	8	9	10	12	13	14	15	16	18	19	20	21
	.05	1	2	3	5	6	7	9	10	12	13	14	16	17	19	20	21	23	24	26
	.10	2	3	5	6	8	9	11	13	14	16	18	19	21	23	24	26	28	29	31
6	.001	0	0	0	0	0	0	2	3	4	5	5	6	7	8	9	10	11	12	13
	.005	0	0	1	2	3	4	5	6	7	8	10	11	12	13	14	16	17	18	19
	.01	0	0	2	3	4	5	7	8	9	10	12	13	14	16	17	19	20	21	23
	.025	0	2	3	4	6	7	9	11	12	14	15	17	18	20	22	23	25	26	28
	.05	1	3	4	6	8	9	11	13	15	17	18	20	22	24	26	27	29	31	33
	.10	2	4	6	8	10	12	14	16	18	20	22	24	26	28	30	32	35	37	39
7	.001	0	0	0	0	1	2	3	4	6	7	8	9	10	11	12	14	15	16	17
	.005	0	0	1	2	4	5	7	8	10	11	13	14	16	17	19	20	22	23	25
	.01	0	1	2	4	5	7	8	10	12	13	15	17	18	20	22	24	25	27	29
	.025	0	2	4	6	7	9	11	13	15	17	19	21	23	25	27	29	31	33	35
	.05	1	3	5	7	9	12	14	16	18	20	22	25	27	29	31	34	36	38	40
	.10	2	5	7	9	12	14	17	19	22	24	27	29	32	34	37	39	42	44	47
8	.001	0	0	0	1	2	3	5	6	7	9	10	12	13	15	16	18	19	21	22
	.005	0	0	2	3	5	7	8	10	12	14	16	18	19	21	23	25	27	29	31
	.01	0	1	3	5	7	8	10	12	14	16	18	21	23	25	27	29	31	33	35
	.025	1	3	5	7	9	11	14	16	18	20	23	25	27	30	32	35	37	39	42
	.05	2	4	6	9	11	14	16	19	21	24	27	29	32	34	37	40	42	45	48
	.10	3	6	8	11	14	17	20	23	25	28	31	34	37	40	43	46	49	52	55

n_1	p	$n_2=2$	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
9	.001	0	0	0	2	3	4	6	8	9	11	13	15	16	18	20	22	24	26	27
	.005	0	1	2	4	6	8	10	12	14	17	19	21	23	25	28	30	32	34	37
	.01	0	2	4	6	8	10	12	15	17	19	22	24	27	29	32	34	37	39	41
	.025	1	3	5	8	11	13	16	18	21	24	27	29	32	35	38	40	43	46	49
	.05	2	5	7	10	13	16	19	22	25	28	31	34	37	40	43	46	49	52	55
	.10	3	6	10	13	16	19	23	26	29	32	36	39	42	46	49	53	56	59	63
10	.001	0	0	1	2	4	6	7	9	11	13	15	18	20	22	24	26	28	30	33
	.005	0	1	3	5	7	10	12	14	17	19	22	25	27	30	32	35	38	40	43
	.01	0	2	4	7	9	12	14	17	20	23	25	28	31	34	37	39	42	45	48
	.025	1	4	6	9	12	15	18	21	24	27	30	34	37	40	43	46	49	53	56
	.05	2	5	8	12	15	18	21	25	28	32	35	38	42	45	49	52	56	59	63
	.10	4	7	11	14	18	22	25	29	33	37	40	44	48	52	55	59	63	67	71
11	.001	0	0	1	3	5	7	9	11	13	16	18	21	23	25	28	30	33	35	38
	.005	0	1	3	6	8	11	14	17	19	22	25	28	31	34	37	40	43	46	49
	.01	0	2	5	8	10	13	16	19	23	26	29	32	35	38	42	45	48	51	54
	.025	1	4	7	10	14	17	20	24	27	31	34	38	41	45	48	52	56	59	63
	.05	2	6	9	13	17	20	24	28	32	35	39	43	47	51	55	58	62	66	70
	.10	4	8	12	16	20	24	28	32	37	41	45	49	53	58	62	66	70	74	79
12	.001	0	0	1	3	5	8	10	13	15	18	21	24	26	29	32	35	38	41	43
	.005	0	2	4	7	10	13	16	19	22	25	28	32	35	38	42	45	48	52	55
	.01	0	3	6	9	12	15	18	22	25	29	32	36	39	43	47	50	54	57	61
	.025	2	5	8	12	15	19	23	27	30	34	38	42	46	50	54	58	62	66	70
	.05	3	6	10	14	18	22	27	31	35	39	43	48	52	56	61	65	69	73	78
	.10	5	9	13	18	22	27	31	36	40	45	50	54	59	64	68	73	78	82	87

n_1	p	$n_2=2$	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
13	.001	0	0	2	4	6	9	12	15	18	21	24	27	30	33	36	39	43	46	49
	.005	0	2	4	8	11	14	18	21	25	28	32	35	39	43	46	50	54	58	61
	.01	1	3	6	10	13	17	21	24	28	32	36	40	44	48	52	56	60	64	68
	.025	2	5	9	13	17	21	25	29	34	38	42	46	51	55	60	64	68	73	77
	.05	3	7	11	16	20	25	29	34	38	43	48	52	57	62	66	71	76	81	85
	.10	5	10	14	19	24	29	34	39	44	49	54	59	64	69	75	80	85	90	95
14	.001	0	0	2	4	7	10	13	16	20	23	26	30	33	37	40	44	47	51	55
	.005	0	2	5	8	12	16	19	23	27	31	35	39	43	47	51	55	59	64	68
	.01	1	3	7	11	14	18	23	27	31	35	39	44	48	52	57	61	66	70	74
	.025	2	6	10	14	18	23	27	32	37	41	46	51	56	60	65	70	75	79	84
	.05	4	8	12	17	22	27	32	37	42	47	52	57	62	67	72	78	83	88	93
	.10	5	11	16	21	26	32	37	42	48	53	59	64	70	75	81	86	92	98	103
15	.001	0	0	2	5	8	11	15	18	22	25	29	33	37	41	44	48	52	56	60
	.005	0	3	6	9	13	17	21	25	30	34	38	43	47	52	56	61	65	70	74
	.01	1	4	8	12	16	20	25	29	34	38	43	48	52	57	62	67	71	76	81
	.025	2	6	11	15	20	25	30	35	40	45	50	55	60	65	71	76	81	86	91
	.05	4	8	13	19	24	29	34	40	45	51	56	62	67	73	78	84	89	95	101
	.10	6	11	17	23	28	34	40	46	52	58	64	69	75	81	87	93	99	105	111
16	.001	0	0	3	6	9	12	16	20	24	28	32	36	40	44	49	53	57	61	66
	.005	0	3	6	10	14	19	23	28	32	37	42	46	51	56	61	66	71	75	80
	.01	1	4	8	13	17	22	27	32	37	42	47	52	57	62	67	72	77	83	88
	.025	2	7	12	16	22	27	32	38	43	48	54	60	65	71	76	82	87	93	99
	.05	4	9	15	20	26	31	37	43	49	55	61	66	72	78	84	90	96	102	108
	.10	6	12	18	24	30	37	43	49	55	62	68	75	81	87	94	100	107	113	120

تابع جدول (8)

n_1	p	$n_2=2$	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
17	.001	0	1	3	6	10	14	18	22	26	30	35	39	44	48	53	58	62	67	71
	.005	0	3	7	11	16	20	25	30	35	40	45	50	55	61	66	71	76	82	87
	.01	1	5	9	14	19	24	29	34	39	45	50	56	61	67	72	78	83	89	94
	.025	3	7	12	18	23	29	35	40	46	52	58	64	70	76	82	88	94	100	106
	.05	4	10	16	21	27	34	40	46	52	58	65	71	78	84	90	97	103	110	116
	.10	7	13	19	26	32	39	46	53	59	66	73	80	86	93	100	107	114	121	128
18	.001	0	1	4	7	11	15	19	24	28	33	38	43	47	52	57	62	67	72	77
	.005	0	3	7	12	17	22	27	32	38	43	48	54	59	65	71	76	82	88	93
	.01	1	5	10	15	20	25	31	37	42	48	54	60	66	71	77	83	89	95	101
	.025	3	8	13	19	25	31	37	43	49	56	62	68	75	81	87	94	100	107	113
	.05	5	10	17	23	29	36	42	49	56	62	69	76	83	89	96	103	110	117	124
	.10	7	14	21	28	35	42	49	56	63	70	78	85	92	99	107	114	121	129	136
19	.001	0	1	4	8	12	16	21	26	30	35	41	46	51	56	61	67	72	78	83
	.005	1	4	8	13	18	23	29	34	40	46	52	58	64	70	75	82	88	94	100
	.01	2	5	10	16	21	27	33	39	45	51	57	64	70	76	83	89	95	102	108
	.025	3	8	14	20	26	33	39	46	53	59	66	73	79	86	93	100	107	114	120
	.05	5	11	18	24	31	38	45	52	59	66	73	81	88	95	102	110	117	124	131
	.10	8	15	22	29	37	44	52	59	67	74	82	90	98	105	113	121	129	136	144
20	.001	0	1	4	8	13	17	22	27	33	38	43	49	55	60	66	71	77	83	89
	.005	1	4	9	14	19	25	31	37	43	49	55	61	68	74	80	87	93	100	106
	.01	2	6	11	17	23	29	35	41	48	54	61	68	74	81	88	94	101	108	115
	.025	3	9	15	21	28	35	42	49	56	63	70	77	84	91	99	106	113	120	128
	.05	5	12	19	26	33	40	48	55	63	70	78	85	93	101	108	116	124	131	139
	.10	8	16	23	31	39	47	55	63	71	79	87	95	103	111	120	128	136	144	152

جدول (9) : مئويات اختبار رتب الإشارة ولكاكسن .

	$W_{0.005}$	$W_{0.01}$	$W_{0.025}$	$W_{0.05}$	$W_{0.10}$	$W_{0.20}$	$W_{0.30}$	$W_{0.40}$	$W_{0.50}$	$\frac{n(n+1)}{2}$
$n=4$	0	0	0	0	1	3	3	4	5	10
5	0	0	0	1	3	4	5	6	7.5	15
6	0	0	1	3	4	6	8	9	10.5	21
7	0	1	3	4	6	9	11	12	14	28
8	1	2	4	6	9	12	14	16	18	36
9	2	4	6	9	11	15	18	20	22.5	45
10	4	6	9	11	15	19	22	25	27.5	55
11	6	8	11	14	18	23	27	30	33	66
12	8	10	14	18	22	28	32	36	39	78
13	10	13	18	22	27	33	38	42	45.5	91
14	13	16	22	26	32	39	44	48	52.5	105
15	16	20	26	31	37	45	51	55	60	120
16	20	24	30	36	43	51	58	63	68	136
17	24	28	35	42	49	58	65	71	76.5	153
18	28	33	41	48	56	66	73	80	85.5	171
19	33	38	47	54	63	74	82	89	95	190
20	38	44	53	61	70	83	91	98	105	210
21	44	50	59	68	78	91	100	108	115.5	131
22	49	56	67	76	87	100	110	119	126.5	153
23	55	63	74	84	95	110	120	130	138	176
24	62	70	82	92	105	120	131	141	150	300
25	69	77	90	101	114	131	143	153	162.5	325
26	76	85	99	111	125	142	155	165	175.5	351
27	84	94	108	120	135	154	167	178	189	378
28	92	102	117	131	146	166	180	192	203	406
29	101	111	127	141	158	178	193	206	217.5	435
30	110	121	138	152	170	191	207	220	232.5	465
31	119	131	148	164	182	205	221	235	248	496
32	129	141	160	176	195	219	236	250	264	528
33	139	152	171	188	208	233	251	266	280.5	561
34	149	163	183	201	222	248	266	282	297.5	595
35	160	175	196	214	236	263	283	299	315	630
36	172	187	209	228	251	279	299	317	333	666
37	184	199	222	242	266	295	316	335	351.5	703
38	196	212	235	257	282	312	334	353	370.5	741
39	208	225	250	272	298	329	352	372	390	780
40	221	239	265	287	314	347	371	391	410	820
41	235	253	280	303	331	365	390	411	430.5	861
42	248	267	295	320	349	384	409	431	451.5	903
43	263	282	311	337	366	403	429	452	473	946
44	277	297	328	354	385	422	450	473	495	990
45	292	313	344	372	403	442	471	495	517.5	1035
46	308	329	362	390	423	463	492	517	540.5	1081
47	324	346	379	408	442	484	514	540	564	1128

المصدر : F. Wilcoxon and R . A .Wilcox , A .C .C . Pearl River , N.Y. ,1964

تابع جدول (9)

	$w_{0.005}$	$w_{0.01}$	$w_{0.025}$	$w_{0.05}$	$w_{0.10}$	$w_{0.20}$	$w_{0.30}$	$w_{0.40}$	$w_{0.50}$	$\frac{n(n+1)}{2}$
48	340	363	397	428	463	505	536	563	588	1176
49	357	381	416	447	483	527	559	587	612.5	1225
50	374	398	435	467	504	550	583	611	637.5	1275

ملحوظة : إذا كانت n أكبر من 50 فإنه يمكن تقريب التجزىء ذو المرتبة α باستخدام العلاقة الآتية : $w_\alpha = \frac{n(n+1)}{4} + z_\alpha \sqrt{n(n+1)(2n+1)/24}$ ، وإذا كانت $\alpha > 0.50$ فإنه يمكن حساب التجزىء بالعلاقة الآتية : $w_\alpha = n(n+1)/2 - w_{1-\alpha}$.

جدول (10) : مئويات اختبار كروسكل – وليس .

حجم العينات	$W_{0.90}$	$W_{0.95}$	$W_{0.99}$
2, 2, 2	3.7143	4.5714	4.5714
3. 2. 1	3.8571	4.2857	4.2857
3. 2. 2	4.4643	4.5000	5.3571
3. 3. 1	4.0000	4.5714	5.1429
3. 3. 2	4.2500	5.1389	6.2500
3. 3. 3	4.6000	5.0667	6.4889
4. 2. 1	4.0179	4.8214	4.8214
4. 2. 2	4.1667	5.1250	6.0000
4. 3. 1	3.8889	5.0000	5.8333
4. 3. 2	4.4444	5.4000	6.3000
4. 3. 3	4.7000	5.7273	6.7091
4, 4. 1	4.0667	4.8667	6.1667
4, 4, 2	4.4455	5.2364	6.8727
4, 4, 3	4.773	5.5758	7.1364
4, 4, 4	4.5000	5.6538	7.5385
5. 2, 1	4.0500	4.4500	5.2500
5. 2. 2	4.2933	5.0400	6.1333
5. 3. 1	3.8400	4.8711	6.4000
5. 3. 2	4.4946	5.1055	6.8218
5, 3, 3	4.4121	5.5152	6.9818
5, 4. 1	3.9600	4.8600	6.8400
5. 4, 2	4.5182	5.2682	7.1182
5, 4. 3	4.5231	5.6308	7.3949
5. 4, 4	4.6187	5.6176	7.7440
5. 5. 1	4.0364	4.9091	6.8364
5, 5, 2	4.5077	5.2462	7.2692
5. 5. 3	4.5363	5.6264	7.5429
5. 5. 4	4.5200	5.6429	7.7914
5. 5. 5	4.5000	5.6600	7.9800

المصدر : Iman , Quade , and Alexander (1975) , American Mathematical Society

جدول (11) : مئويات اختبار كولو مجروف – سميئروف .

اختبار من طرف واحد	$p = .90$	.95	.975	.99	.995		$p = .90$	.95	.975	.99	.995
اختبار من طرفين	$p = .80$	.90	.95	.98	.99		$p = .80$	.90	.95	.98	.99
$n = 1$	.900	.950	.975	.990	.995	$n = 21$	.226	.259	.287	.321	.344
2	.684	.776	.842	.900	.929	22	.221	.253	.281	.314	.337
3	.565	.636	.708	.785	.829	23	.216	.247	.275	.307	.330
4	.493	.565	.624	.689	.734	24	.212	.242	.269	.301	.323
5	.447	.509	.563	.627	.669	25	.208	.238	.264	.295	.317
6	.410	.468	.519	.577	.617	26	.204	.233	.259	.290	.311
7	.381	.436	.483	.538	.576	27	.200	.229	.254	.284	.305
8	.358	.410	.454	.507	.542	28	.197	.225	.250	.279	.300
9	.339	.387	.430	.480	.513	29	.193	.221	.246	.275	.295
10	.323	.369	.409	.457	.489	30	.190	.218	.242	.270	.290
11	.308	.352	.391	.437	.468	31	.187	.214	.238	.266	.285
12	.296	.338	.375	.419	.449	32	.184	.211	.234	.262	.281
13	.285	.325	.361	.404	.432	33	.182	.208	.231	.258	.277
14	.275	.314	.349	.390	.418	34	.179	.205	.227	.254	.273
15	.266	.304	.338	.377	.404	35	.177	.202	.224	.251	.269
16	.258	.295	.327	.366	.392	36	.174	.199	.221	.247	.265
17	.250	.286	.318	.355	.381	37	.172	.196	.218	.244	.262
18	.244	.279	.309	.346	.371	38	.170	.194	.215	.241	.258
19	.237	.271	.301	.337	.361	39	.168	.191	.213	.238	.255
20	.232	.265	.294	.329	.352	40	.165	.189	.210	.235	.252
						Approximation for $n > 40$	$\frac{1.07}{\sqrt{n}}$	$\frac{1.22}{\sqrt{n}}$	$\frac{1.36}{\sqrt{n}}$	$\frac{1.52}{\sqrt{n}}$	$\frac{1.63}{\sqrt{n}}$

المصدر :

L. H. Miller , " Tables of Percentage Points of Kolmogorov Statistics ," JASA. ,51 (1956) , 111 - 121

جدول (12) : مئويات اختبار سبيرمان .

n	p = .900	.950	.975	.990	.995	.999
4	.8000	.8000				
5	.7000	.8000	.9000	.9000		
6	.6000	.7714	.8286	.8857	.9429	
7	.5357	.6786	.7450	.8571	.8929	.9643
8	.5000	.6190	.7143	.8095	.8571	.9286
9	.4667	.5833	.6833	.7667	.8167	.9000
10	.4424	.5515	.6364	.7333	.7818	.8667
11	.4182	.5273	.6091	.7000	.7455	.8364
12	.3986	.4965	.5804	.6713	.7273	.8182
13	.3791	.4780	.5549	.6429	.6978	.7912
14	.3626	.4593	.5341	.6220	.6747	.7670
15	.3500	.4429	.5179	.6000	.6536	.7464
16	.3382	.4265	.5000	.5824	.6324	.7265
17	.3260	.4118	.4853	.5637	.6152	.7083
18	.3148	.3994	.4716	.5480	.5975	.6904
19	.3070	.3895	.4579	.5333	.5825	.6737
20	.2977	.3789	.4451	.5203	.5684	.6586
21	.2909	.3688	.4351	.5078	.5545	.6455
22	.2829	.3597	.4241	.4963	.5426	.6318
23	.2767	.3518	.4150	.4852	.5306	.6186
24	.2704	.3435	.4061	.4748	.5200	.6070
25	.2646	.3362	.3977	.4654	.5100	.5962
26	.2588	.3299	.3894	.4564	.5002	.5856
27	.2540	.3236	.3822	.4481	.4915	.5757
28	.2490	.3175	.3749	.4401	.4828	.5660
29	.2443	.3113	.3685	.4320	.4744	.5567
30	.2400	.3059	.3620	.4251	.4665	.5479

المصدر : Glasser and Winter (1961) , Biometrika Trustees .

جدول (13) : التحويل من r إلى z .

الجدول يتضمن قيم $z = 0.5[\ln(1+r)/(1-r)] = \tanh^{-1} r$ المقابلة لقيم r .

r	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.00000	.01000	.02000	.03001	.04002	.05004	.06007	.07012	.08017	.09024
.1	.10034	.11045	.12058	.13074	.14093	.15114	.16139	.17167	.18198	.19234
.2	.20273	.21317	.22366	.23419	.24477	.25541	.26611	.27686	.28768	.29857
.3	.30952	.32055	.33165	.34283	.35409	.36344	.37689	.38842	.40006	.41180
.4	.42365	.43561	.44769	.45990	.47223	.48470	.49731	.51007	.52298	.53606
.5	.54931	.56273	.57634	.59014	.60415	.61838	.63283	.64752	.66246	.67767
.6	.69315	.70892	.72500	.74142	.75817	.77530	.79281	.81074	.82911	.84795
.7	.86730	.88718	.90764	.92873	.95048	.97295	.99621	1.02033	1.04537	1.07143
.8	1.09861	1.12703	1.15682	1.18813	1.22117	1.25615	1.29334	1.33308	1.37577	1.42192
.9	1.47222	1.52752	1.58902	1.65839	1.73805	1.83178	1.94591	2.09229	2.29756	2.64663

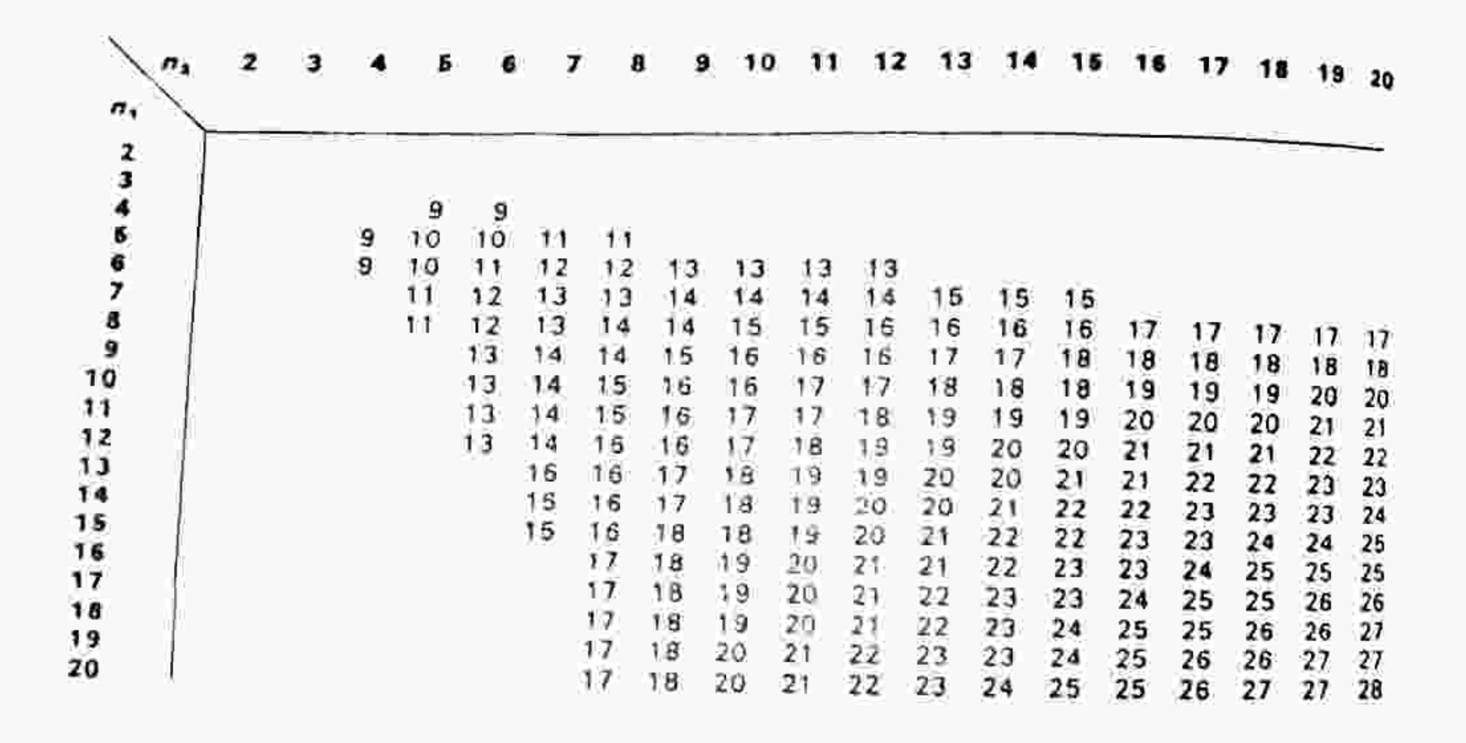
المصدر :

Introduction to Statistical Analysis , 3rd. ed. , by W . J . Dixon and F . J . Massey , 1969 by McGraw-Hill , Inc.

جدول (14) : القيم الجدولية لاختبار العشوائية .

قيم r السفلى الحرجة لاختبار العشوائية عندما $\alpha = 0.05$:

n_1 \ n_2	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
2											2	2	2	2	2	2	2	2	2
3					2	2	2	2	2	2	2	2	2	3	3	3	3	3	3
4				2	2	2	3	3	3	3	3	3	3	3	4	4	4	4	4
5			2	2	3	3	3	3	3	4	4	4	4	4	4	4	5	5	5
6		2	2	3	3	3	3	4	4	4	4	5	5	5	5	5	5	6	6
7		2	2	3	3	3	4	4	5	5	5	5	5	6	6	6	6	6	6
8		2	3	3	3	4	4	5	5	5	6	6	6	6	6	7	7	7	7
9		2	3	3	4	4	5	5	5	6	6	6	7	7	7	7	8	8	8
10		2	3	3	4	5	5	5	6	6	7	7	7	7	8	8	8	8	9
11		2	3	4	4	5	5	6	6	7	7	7	8	8	8	9	9	9	9
12	2	2	3	4	4	5	6	6	7	7	7	8	8	8	9	9	9	10	10
13	2	2	3	4	5	5	6	6	7	7	8	8	9	9	9	10	10	10	10
14	2	2	3	4	5	5	6	7	7	8	8	9	9	9	10	10	10	11	11
15	2	3	3	4	5	6	6	7	7	8	8	9	9	10	10	11	11	11	12
16	2	3	4	4	5	6	6	7	8	8	9	9	10	10	11	11	11	12	12
17	2	3	4	4	5	6	7	7	8	9	9	10	10	11	11	11	12	12	13
18	2	3	4	5	5	6	7	8	8	9	9	10	10	11	11	12	12	13	13
19	2	3	4	5	6	6	7	8	8	9	10	10	11	11	12	12	13	13	13
20	2	3	4	5	6	6	7	8	9	9	10	10	11	12	12	13	13	13	14



قيم r العليا الحرجة لاختبار العشوائية عندما $\alpha = 0.05$:

n_1 \ n_2	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
2																			
3																			
4				9	9														
5			9	10	10	11	11												
6			9	10	11	12	12	13	13	13	13								
7				11	12	13	13	14	14	14	14	15	15	15					
8				11	12	13	14	14	15	15	16	16	16	16	17	17	17	17	17
9					13	14	14	15	16	16	16	17	17	18	18	18	18	18	18
10					13	14	15	16	16	17	17	18	18	18	19	19	19	20	20
11					13	14	15	16	17	17	18	19	19	19	20	20	20	21	21
12					13	14	15	16	17	18	19	19	20	20	21	21	21	22	22
13						16	16	17	18	19	19	20	20	21	21	22	22	23	23
14						15	16	17	18	19	20	20	21	22	22	23	23	23	24
15						15	16	18	18	19	20	21	22	22	23	23	24	24	25
16							17	18	19	20	21	21	22	23	23	24	25	25	25
17							17	18	19	20	21	22	23	23	24	25	25	26	26
18							17	18	19	20	21	22	23	24	25	25	26	26	27
19							17	18	20	21	22	23	23	24	25	26	26	27	27
20							17	18	20	21	22	23	24	25	25	26	27	27	28

المصدر :

Frieda S. Swed and C. Eisenhart , " Tables for Testing Randomness of Grouping in a Sequence of Alternatives , " Ann. Math. Statist. , 14 (1943) , 66-87

جدول (15) : القيم الجدولية لاختبار دونيت : من طرفين عندما $\alpha = 0.05$.

N - t	عدد المعالجات باستثناء معالجة المراقبة $(t-1)$								
	1	2	3	4	5	6	7	8	9
5	2.57	3.03	3.29	3.48	3.62	3.73	3.82	3.90	3.97
6	2.45	2.86	3.10	3.26	3.39	3.49	3.57	3.64	3.71
7	2.36	2.75	2.97	3.12	3.24	3.33	3.41	3.47	3.53
8	2.31	2.67	2.88	3.02	3.13	3.22	3.29	3.35	3.41
9	2.26	2.61	2.81	2.95	3.05	3.14	3.20	3.26	3.32
10	2.23	2.57	2.76	2.89	2.99	3.07	3.14	3.19	3.24
11	2.20	2.53	2.72	2.84	2.94	3.02	3.08	3.14	3.19
12	2.18	2.50	2.68	2.81	2.90	2.98	3.04	3.09	3.14
13	2.16	2.48	2.65	2.78	2.87	2.94	3.00	3.06	3.10
14	2.14	2.46	2.63	2.75	2.84	2.91	2.97	3.02	3.07
15	2.13	2.44	2.61	2.73	2.82	2.89	2.95	3.00	3.04
16	2.12	2.42	2.59	2.71	2.80	2.87	2.92	2.97	3.02
17	2.11	2.41	2.58	2.69	2.78	2.85	2.90	2.95	3.00
18	2.10	2.40	2.56	2.68	2.76	2.83	2.89	2.94	2.98
19	2.09	2.39	2.55	2.66	2.75	2.81	2.87	2.92	2.96
20	2.09	2.38	2.54	2.65	2.73	2.80	2.86	2.90	2.95
24	2.06	2.35	2.51	2.61	2.70	2.76	2.81	2.86	2.90
30	2.04	2.32	2.47	2.58	2.66	2.72	2.77	2.82	2.86
40	2.02	2.29	2.44	2.54	2.62	2.68	2.73	2.77	2.81
60	2.00	2.27	2.41	2.51	2.58	2.64	2.69	2.73	2.77
120	1.98	2.24	2.38	2.47	2.55	2.60	2.65	2.69	2.73
∞	1.96	2.21	2.35	2.44	2.51	2.57	2.61	2.65	2.69

المصدر :

C . W . Dunnett , " New Tables for Multiple Comparison with a Control . " Biometrics, vol. 20 , no. 3, 1964 , and from JASA , vol. 50, 1955 .

تابع جدول (15) : من طرف واحد عندما $\alpha = 0.05$.

N - t	عدد المعالجات باستثناء معالجة المراقبة $(t - 1)$								
	1	2	3	4	5	6	7	8	9
5	2.02	2.44	2.68	2.85	2.98	3.08	3.16	3.24	3.30
6	1.94	2.34	2.56	2.71	2.83	2.92	3.00	3.07	3.12
7	1.89	2.27	2.48	2.62	2.73	2.82	2.89	2.95	3.01
8	1.86	2.22	2.42	2.55	2.66	2.74	2.81	2.87	2.92
9	1.83	2.18	2.37	2.50	2.60	2.68	2.75	2.81	2.86
10	1.81	2.15	2.34	2.47	2.56	2.64	2.70	2.76	2.81
11	1.80	2.13	2.31	2.44	2.53	2.60	2.67	2.72	2.77
12	1.78	2.11	2.29	2.41	2.50	2.58	2.64	2.69	2.74
13	1.77	2.09	2.27	2.39	2.48	2.55	2.61	2.66	2.71
14	1.76	2.08	2.25	2.37	2.46	2.53	2.59	2.64	2.69
15	1.75	2.07	2.24	2.36	2.44	2.51	2.57	2.62	2.67
16	1.75	2.06	2.23	2.34	2.43	2.50	2.56	2.61	2.65
17	1.74	2.05	2.22	2.33	2.42	2.49	2.54	2.59	2.64
18	1.73	2.04	2.21	2.32	2.41	2.48	2.53	2.58	2.62
19	1.73	2.03	2.20	2.31	2.40	2.47	2.52	2.57	2.61
20	1.72	2.03	2.19	2.30	2.39	2.46	2.51	2.56	2.60
24	1.71	2.01	2.17	2.28	2.36	2.43	2.48	2.53	2.57
30	1.70	1.99	2.15	2.25	2.33	2.40	2.45	2.50	2.54
40	1.68	1.97	2.13	2.23	2.31	2.37	2.42	2.47	2.51
60	1.67	1.95	2.10	2.21	2.28	2.35	2.39	2.44	2.48
120	1.66	1.93	2.08	2.18	2.26	2.32	2.37	2.41	2.45
∞	1.64	1.92	2.06	2.16	2.23	2.29	2.34	2.38	2.42

تابع جدول (15) : من طرفين عندما $\alpha = 0.01$.

N - t	عدد المعالجات باستثناء معالجة المراقبة $(t-1)$								
	1	2	3	4	5	6	7	8	9
5	4.03	4.63	4.98	5.22	5.41	5.56	5.69	5.80	5.89
6	3.71	4.21	4.51	4.71	4.87	5.00	5.10	5.20	5.28
7	3.50	3.95	4.21	4.39	4.53	4.64	4.74	4.82	4.89
8	3.36	3.77	4.00	4.17	4.29	4.40	4.48	4.56	4.62
9	3.25	3.63	3.85	4.01	4.12	4.22	4.30	4.37	4.43
10	3.17	3.53	3.74	3.88	3.99	4.08	4.16	4.22	4.28
11	3.11	3.45	3.65	3.79	3.89	3.98	4.05	4.11	4.16
12	3.05	3.39	3.58	3.71	3.81	3.89	3.96	4.02	4.07
13	3.01	3.33	3.52	3.65	3.74	3.82	3.89	3.94	3.99
14	2.98	3.29	3.47	3.59	3.69	3.76	3.83	3.88	3.93
15	2.95	3.25	3.43	3.55	3.64	3.71	3.78	3.83	3.88
16	2.92	3.22	3.39	3.51	3.60	3.67	3.73	3.78	3.83
17	2.90	3.19	3.36	3.47	3.56	3.63	3.69	3.74	3.79
18	2.88	3.17	3.33	3.44	3.53	3.60	3.66	3.71	3.75
19	2.86	3.15	3.31	3.42	3.50	3.57	3.63	3.68	3.72
20	2.85	3.13	3.29	3.40	3.48	3.55	3.60	3.65	3.69
24	2.80	3.07	3.22	3.32	3.40	3.47	3.52	3.57	3.61
30	2.75	3.01	3.15	3.25	3.33	3.39	3.44	3.49	3.52
40	2.70	2.95	3.09	3.19	3.26	3.32	3.37	3.41	3.44
60	2.66	2.90	3.03	3.12	3.19	3.25	3.29	3.33	3.37
120	2.62	2.85	2.97	3.06	3.12	3.18	3.22	3.26	3.29
∞	2.58	2.79	2.92	3.00	3.06	3.11	3.15	3.19	3.22

تابع جدول (15) : من طرف واحد عندما $\alpha = 0.01$.

N - t	عدد المعالجات باستثناء معالجة المراقبة $(t-1)$								
	1	2	3	4	5	6	7	8	9
5	4.03	4.63	4.98	5.22	5.41	5.56	5.69	5.80	5.89
6	3.71	4.21	4.51	4.71	4.87	5.00	5.10	5.20	5.28
7	3.50	3.95	4.21	4.39	4.53	4.64	4.74	4.82	4.89
8	3.36	3.77	4.00	4.17	4.29	4.40	4.48	4.56	4.62
9	3.25	3.63	3.85	4.01	4.12	4.22	4.30	4.37	4.43
10	3.17	3.53	3.74	3.88	3.99	4.08	4.16	4.22	4.28
11	3.11	3.45	3.65	3.79	3.89	3.98	4.05	4.11	4.16
12	3.05	3.39	3.58	3.71	3.81	3.89	3.96	4.02	4.07
13	3.01	3.33	3.52	3.65	3.74	3.82	3.89	3.94	3.99
14	2.98	3.29	3.47	3.59	3.69	3.76	3.83	3.88	3.93
15	2.95	3.25	3.43	3.55	3.64	3.71	3.78	3.83	3.88
16	2.92	3.22	3.39	3.51	3.60	3.67	3.73	3.78	3.83
17	2.90	3.19	3.36	3.47	3.56	3.63	3.69	3.74	3.79
18	2.88	3.17	3.33	3.44	3.53	3.60	3.66	3.71	3.75
19	2.86	3.15	3.31	3.42	3.50	3.57	3.63	3.68	3.72
20	2.85	3.13	3.29	3.40	3.48	3.55	3.60	3.65	3.69
24	2.80	3.07	3.22	3.32	3.40	3.47	3.52	3.57	3.61
30	2.75	3.01	3.15	3.25	3.33	3.39	3.44	3.49	3.52
40	2.70	2.95	3.09	3.19	3.26	3.32	3.37	3.41	3.44
60	2.66	2.90	3.03	3.12	3.19	3.25	3.29	3.33	3.37
120	2.62	2.85	2.97	3.06	3.12	3.18	3.22	3.26	3.29
∞	2.58	2.79	2.92	3.00	3.06	3.11	3.15	3.19	3.22

جدول (16) : القيم المئوية لاختبار توكي عندما $\alpha = 0.05$.

Error df	2	3	4	5	6	7	8	9	10
1	17.97	26.98	32.82	37.08	40.41	43.12	45.40	47.36	49.07
2	6.08	8.33	9.80	10.88	11.74	12.44	13.03	13.54	13.99
3	4.50	5.91	6.82	7.50	8.04	8.48	8.85	9.18	9.46
4	3.93	5.04	5.76	6.29	6.71	7.05	7.35	7.60	7.83
5	3.64	4.60	5.22	5.67	6.03	6.33	6.58	6.80	6.99
6	3.46	4.34	4.90	5.30	5.63	5.90	6.12	6.32	6.49
7	3.34	4.16	4.68	5.06	5.36	5.61	5.82	6.00	6.16
8	3.26	4.04	4.53	4.89	5.17	5.40	5.60	5.77	5.92
9	3.20	3.95	4.41	4.76	5.02	5.24	5.43	5.59	5.74
10	3.15	3.88	4.33	4.65	4.91	5.12	5.30	5.46	5.60
11	3.11	3.82	4.26	4.57	4.82	5.03	5.20	5.35	5.49
12	3.08	3.77	4.20	4.51	4.75	4.95	5.12	5.27	5.39
13	3.06	3.73	4.15	4.45	4.69	4.88	5.05	5.19	5.32
14	3.03	3.70	4.11	4.41	4.64	4.83	4.99	5.13	5.25
15	3.01	3.67	4.08	4.37	4.59	4.78	4.94	5.08	5.20
16	3.00	3.65	4.05	4.33	4.56	4.74	4.90	5.03	5.15
17	2.98	3.63	4.02	4.30	4.52	4.70	4.86	4.99	5.11
18	2.97	3.61	4.00	4.28	4.49	4.67	4.82	4.96	5.07
19	2.96	3.59	3.98	4.25	4.47	4.65	4.79	4.92	5.04
20	2.95	3.58	3.96	4.23	4.45	4.62	4.77	4.90	5.01
24	2.92	3.53	3.90	4.17	4.37	4.54	4.68	4.81	4.92
30	2.89	3.49	3.85	4.10	4.30	4.46	4.60	4.72	4.82
40	2.86	3.44	3.79	4.04	4.23	4.39	4.52	4.63	4.73
60	2.83	3.40	3.74	3.98	4.16	4.31	4.44	4.55	4.65
120	2.80	3.36	3.68	3.92	4.10	4.24	4.36	4.47	4.56
∞	2.77	3.31	3.63	3.86	4.03	4.17	4.29	4.39	4.47

Error df	11	12	13	14	15	16	17	18	19	20
1	50.59	51.96	53.20	54.33	55.36	56.32	57.22	58.04	58.83	59.56
2	14.39	14.75	15.08	15.38	15.65	15.91	16.14	16.37	16.57	16.77
3	9.72	9.95	10.15	10.35	10.52	10.69	10.84	10.98	11.11	11.24
4	8.03	8.21	8.37	8.52	8.66	8.79	8.91	9.03	9.13	9.23
5	7.17	7.32	7.47	7.60	7.72	7.83	7.93	8.03	8.12	8.21
6	6.65	6.79	6.92	7.03	7.14	7.24	7.34	7.43	7.51	7.59
7	6.30	6.43	6.55	6.66	6.76	6.85	6.94	7.02	7.10	7.17
8	6.05	6.18	6.29	6.39	6.48	6.57	6.65	6.73	6.80	6.87
9	5.87	5.98	6.09	6.19	6.28	6.36	6.44	6.51	6.58	6.64
10	5.72	5.83	5.93	6.03	6.11	6.19	6.27	6.34	6.40	6.47
11	5.61	5.71	5.81	5.90	5.98	6.06	6.13	6.20	6.27	6.33
12	5.51	5.61	5.71	5.80	5.88	5.95	6.02	6.09	6.15	6.21
13	5.43	5.53	5.63	5.71	5.79	5.86	5.93	5.99	6.05	6.11
14	5.36	5.46	5.55	5.64	5.71	5.79	5.85	5.91	5.97	6.03
15	5.31	5.40	5.49	5.57	5.65	5.72	5.78	5.85	5.90	5.96
16	5.26	5.35	5.44	5.52	5.59	5.66	5.73	5.79	5.84	5.90
17	5.21	5.31	5.39	5.47	5.54	5.61	5.67	5.73	5.79	5.84
18	5.17	5.27	5.35	5.43	5.50	5.57	5.63	5.69	5.74	5.79
19	5.14	5.23	5.31	5.39	5.46	5.53	5.59	5.65	5.70	5.75
20	5.11	5.20	5.28	5.36	5.43	5.49	5.55	5.61	5.66	5.71
24	5.01	5.10	5.18	5.25	5.32	5.38	5.44	5.49	5.55	5.59
30	4.92	5.00	5.08	5.15	5.21	5.27	5.33	5.38	5.43	5.47
40	4.82	4.90	4.98	5.04	5.11	5.16	5.22	5.27	5.31	5.36
60	4.73	4.81	4.88	4.94	5.00	5.06	5.11	5.15	5.20	5.24
120	4.64	4.71	4.78	4.84	4.90	4.95	5.00	5.04	5.09	5.13
∞	4.55	4.62	4.68	4.74	4.80	4.85	4.89	4.93	4.97	5.01

المصدر : D . B . Duncan , " Multiple range and multiple F tests , " Biometrics ,11: 1-42 (1955) .

تابع جدول (16) : عندما $\alpha = 0.01$.

Error df	2	3	4	5	6	7	8	9	10
1	90.03	135.0	164.3	185.6	202.2	215.8	227.2	237.0	245.6
2	14.04	19.02	22.29	24.72	26.63	28.20	29.53	30.68	31.69
3	8.26	10.62	12.17	13.33	14.24	15.00	15.64	16.20	16.69
4	6.51	8.12	9.17	9.96	10.58	11.10	11.55	11.93	12.27
5	5.70	6.98	7.80	8.42	8.91	9.32	9.67	9.97	10.24
6	5.24	6.33	7.03	7.56	7.97	8.32	8.61	8.87	9.10
7	4.95	5.92	6.54	7.01	7.37	7.68	7.94	8.17	8.37
8	4.75	5.64	6.20	6.62	6.96	7.24	7.47	7.68	7.86
9	4.60	5.43	5.96	6.35	6.66	6.91	7.13	7.33	7.49
10	4.48	5.27	5.77	6.14	6.43	6.67	6.87	7.05	7.21
11	4.39	5.15	5.62	5.97	6.25	6.48	6.67	6.84	6.99
12	4.32	5.05	5.50	5.84	6.10	6.32	6.51	6.67	6.81
13	4.26	4.96	5.40	5.73	5.98	6.19	6.37	6.53	6.67
14	4.21	4.89	5.32	5.63	5.88	6.08	6.26	6.41	6.54
15	4.17	4.84	5.25	5.56	5.80	5.99	6.16	6.31	6.44
16	4.13	4.79	5.19	5.49	5.72	5.92	6.08	6.22	6.35
17	4.10	4.74	5.14	5.43	5.66	5.85	6.01	6.15	6.27
18	4.07	4.70	5.09	5.38	5.60	5.79	5.94	6.08	6.20
19	4.05	4.67	5.05	5.33	5.55	5.73	5.89	6.02	6.14
20	4.02	4.64	5.02	5.29	5.51	5.69	5.84	5.97	6.09
24	3.96	4.55	4.91	5.17	5.37	5.54	5.69	5.81	5.92
30	3.89	4.45	4.80	5.05	5.24	5.40	5.54	5.65	5.76
40	3.82	4.37	4.70	4.93	5.11	5.26	5.39	5.50	5.60
60	3.76	4.28	4.59	4.82	4.99	5.13	5.25	5.36	5.45
120	3.70	4.20	4.50	4.71	4.87	5.01	5.12	5.21	5.30
∞	3.64	4.12	4.40	4.60	4.76	4.88	4.99	5.08	5.16

Error df	11	12	13	14	15	16	17	18	19	20
1	253.2	260.0	266.2	271.8	277.0	281.8	286.3	290.4	294.3	298.0
2	32.59	33.40	34.13	34.81	35.43	36.00	36.53	37.03	37.50	37.95
3	17.13	17.53	17.89	18.22	18.52	18.81	19.07	19.32	19.55	19.77
4	12.57	12.84	13.09	13.32	13.53	13.73	13.91	14.08	14.24	14.40
5	10.48	10.70	10.89	11.08	11.24	11.40	11.55	11.68	11.81	11.93
6	9.30	9.48	9.65	9.81	9.95	10.08	10.21	10.32	10.43	10.54
7	8.55	8.71	8.86	9.00	9.12	9.24	9.35	9.46	9.55	9.65
8	8.03	8.18	8.31	8.44	8.55	8.66	8.76	8.85	8.94	9.03
9	7.65	7.78	7.91	8.03	8.13	8.23	8.33	8.41	8.49	8.57
10	7.36	7.49	7.60	7.71	7.81	7.91	7.99	8.08	8.15	8.23
11	7.13	7.25	7.36	7.46	7.56	7.65	7.73	7.81	7.88	7.95
12	6.94	7.06	7.17	7.26	7.36	7.44	7.52	7.59	7.66	7.73
13	6.79	6.90	7.01	7.10	7.19	7.27	7.35	7.42	7.48	7.55
14	6.66	6.77	6.87	6.96	7.05	7.13	7.20	7.27	7.33	7.39
15	6.55	6.66	6.76	6.84	6.93	7.00	7.07	7.14	7.20	7.26
16	6.46	6.56	6.66	6.74	6.82	6.90	6.97	7.03	7.09	7.15
17	6.38	6.48	6.57	6.66	6.73	6.81	6.87	6.94	7.00	7.05
18	6.31	6.41	6.50	6.58	6.65	6.73	6.79	6.85	6.91	6.97
19	6.25	6.34	6.43	6.51	6.58	6.65	6.72	6.78	6.84	6.89
20	6.19	6.28	6.37	6.45	6.52	6.59	6.65	6.71	6.77	6.82
24	6.02	6.11	6.19	6.26	6.33	6.39	6.45	6.51	6.56	6.61
30	5.85	5.93	6.01	6.08	6.14	6.20	6.26	6.31	6.36	6.41
40	5.69	5.76	5.83	5.90	5.96	6.02	6.07	6.12	6.16	6.21
60	5.53	5.60	5.67	5.73	5.78	5.84	5.89	5.93	5.97	6.01
120	5.37	5.44	5.50	5.56	5.61	5.66	5.71	5.75	5.79	5.83
∞	5.23	5.29	5.35	5.40	5.45	5.49	5.54	5.57	5.61	5.65

قائمة المراجع :

Conover , W. J. , Practical Nonparametric Statistics , 2nd ed. John-Wiley and Sons , New York , 1980 .

Daniel , W. W., Applied Nonparametric Statistics , Houghton Mifflin Co. , New Jersey , 1978 .

Daniel ,W. W., Biostatistics Afoundation for Analysis in the Health Sciences , 3rd ed. John-Wiley and Sons ,New York , 1976 .

Degroot ,V. Probability and Statistics ,Addison - Welsy Publishing Co. , California , 1975 .

Dixon , W . J . and Massey , F . J . , Introduction to Statistical Analysis , 3rd. ed. McGraw-Hill , Inc. ,1969 .

Draper , N. R. and Smith , H. , Applied Regression Analysis , 2nd ed. , Wiley , New York , 1981 .

Dwass , M. , First Steps in Probability , Mcgraw-Hill , New York ,1967 .

Fakki , H., Method for Fram Surveys and On-Fram Trials ,ICARDA , Syria ,1989.

Feller ,W. , An Introduction to Probabilty Theory and Its Applications ,Vol.II , John Wiley, New york ,1966 .

Goon , A. ,Gupta ,K. and Dasgupta ,B. , An OutLine of Statistical Theory ,Vol.I.,The World Press Ltd. ,Calcutta,1970.

Hajek , J. ,Theory of Rank Tests , Academic Press ,New York ,1967.

Hogg ,R.V. , and Craig , A.T., Introduction to Mathematical Statistics ,3rd ed. The Macmillan Co. , New York ,1970.

John , P. M. , Statistical Design and Analysis of Experiments , Macmillan , New York , 1971 .

Lindgren , B. W. , Statistical Thoery , 2nd ed. The Macmillan Co. ,New York ,1968 .

Madsen ,R. W. and Moeschberger , M. L. Statistical Concepts With Applications to Business and Economics , Prentice-Hall , New Jersey , 1980 .

Malik ,H. J, and Mullen ,K. ,A First Course In Probability and Statistics , AddisonWesley Publishing Co. ,California ,1973.

McClave , J. T. and Dietrich , F. H. Statistics , Dellen Co. , California , 1979 .

Mendenhall , W. , Scheaffer , L. , and Wackerly , D. Mathematical Statistics with Applications , 2nd ed. , Dusbury Press , California , 1981 .

Mood , A.M., and Graybill ,F.A. ,Introduction to The Theory of Statistics , 2nd ed., McGraw - Hill ,New York ,1963 .

Rohatgi ,V., An Introduction to Probability Theory and Mathematical Statistics , John-Wiley and Sons , New York , 1976.

Ross , S., A First course in Probability ,Macmillan Publishing Co.,Inc. 1976 .

Roussas ,G., A First Course in Mathematical Statistics ,Addison-Wesley Publishing Co., California , 1973.

Smillie , K. W. , An Introduction to Regression and Correlation , Academic Press , New York , 1966 .

Walpole ,R. E. and Myers , R.H. , Probability and Statistics for Enginers and Scientists, Macmilan Publishing Co., New York ,1985.